Resonance

December 1996 Volume 1 Number 12

journal of science education

Selberg on Ramanujan ❖
Can We Pop a Pill to Cure Obesity? ❖
Statistical Thermodynamics ❖
Approach to Absolute Zero ❖ Algorithms

Resonance **is a monthly journal of science education published by Indian Academy of Sciences, Bangalore, India.**

Editorial Office: Indian Academy of Sciences, C V Raman Avenue, PB No. 8005, Bangalore 560 080, India.
Tel:+91 (80) 3342310 / 3342546, Fax:+91 (80) 3346094, email:resonanc@ias.ernet.in

Editorial

N Mukunda, Chief Editor

With this issue we complete the first year and volume of *Resonance.* At the same time the present editorial team prepares to hand over to a new one which will take charge from the New Year. For all of us it has been a rich and educative experience. Many letters have come to us, some appreciative and others critical. The former are pleasant to read, but in some ways the latter are more important. One thing we have learnt – the gift of writing expository articles on science and mathematics, accessible to a wide audience, is quite rare and also difficult to cultivate. For many reasons, this seems to be particularly so in our country. But we have also been encouraged to see some really talented writers as well. As editors we have had to judge material in our hands, even when solicited by us, as objectively as possible; though the final decisions, reflecting our own backgrounds and judgements, may sometimes be viewed as subjective! We hope the new editorial team will build upon this year's efforts and, given the unavoidable constraints, continually improve the quality and appeal of *Resonance* .

The twentysecond of December is the date of birth of Srinivasa Ramanujan Iyengar. In 1987 his birth centenary was celebrated the world over, and Robert Kanigel's marvellous biography "The man who knew infinity" appeared just around then. We feature Ramanujan in several ways in this issue: the familiar haunting passport photo on the back cover; an account by Shailesh Shirali of a problem on continued fractions to which Ramanujan had contributed an answer ; a review of Kanigel's book by Rajat Tandon; another by Vittal Rao of G H Hardy's "A Mathematician's Apology"; and finally the text of an extempore lecture to a

The gift of writing expository articles on science and mathematics, accessible to a wide audience, is quite rare and also difficult to cultivate.

Only the universality of mathematics can explain how the Englishman Hardy and the South Indian Ramanujan could communicate so freely in technical matters, even while being unable to cross cultural barriers.

general audience given by Atle Selberg at the Tata Institute of Fundamental Research, Mumbai, on the occasion of the centenary celebrations in 1987.

As Tandon remarks, only the universality of mathematics can explain how the Englishman Hardy and the South Indian Ramanujan could communicate so freely in technical matters, even while being unable to cross cultural barriers. And in the same vein we may say that is also how Ramanujan could end up deeply influencing Selberg in "remote Norway". Fortunately both Kanigel's book and Hardy's "Apology" are available at reasonable prices; and the latter has an introduction by C P Snow dealing largely with the Ramanujan-Hardy story. There are attempts now and again to trace the roots of Ramanujan's genius, his methods of working and presentation of results, to earlier existing traditions in Indian mathematics. In this context, Selberg's remark about Evariste Galois in France and Niels Henrik Abel in Norway (who both died very young) - "although they had their difficulties, come from a somewhat more fortunate environment" - is very significant. To my mind, the answer to the question in what sense Ramanujan was a child and product of India, goes beyond mathematics and covers all forms of such genius; and is perhaps most poetically captured by Kahlil Gibran's lines in "The Prophet":

> *Your children are not your children.*
> *They are the sons and daughters of life's longing for itself.*
> *They come through you but not from you,*
> *And though they are with you yet they belong not to you.*
> *You may give them your love but not your thoughts,*
> *For they have their own thoughts.*
> *You may house their bodies but not their souls,*
> *For their souls dwell in the house of tomorrow,*
> *which you cannot visit, not even in your dreams.*

Science Smiles

R K Laxman

Thanks doctor, I am able to see far better now.

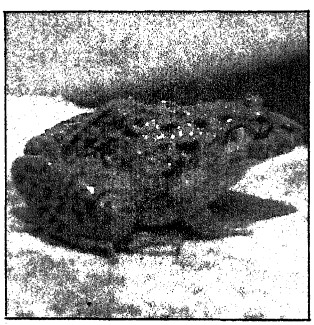
Resonance journal of science education

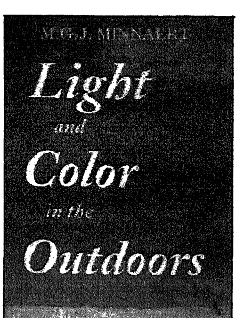
December 1996 Volume 1 Number 12

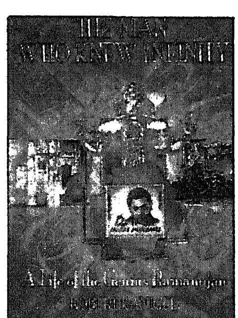
SERIES ARTICLES

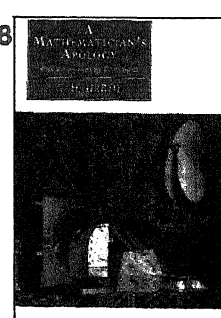
17 

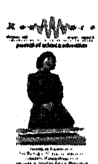
GENERAL ARTICLES

39 

73 

76 

78 



Front Cover

SRL Levitates Woman 5cm with Single Crystal YBCO. Researchers at the International Superconductivity Technology Center's Superconductivity Research Laboratory (SRL) reported that they have levitated a woman (Ms Asakawa) with 5 centimeters clearance using (HTS) materials, beating the old record (5 millimeters) a factor of ten.

Back Cover

Srinivasa Ramanujan (December 1887 - April 1920)

Approach To Absolute Zero

1. Liquefaction Of Gases

R Srinivasan

R Srinivasan is a Visiting Professor at the Raman Research Institute after retiring as Director of the Inter-University Consortium for DAE Facilities at Indore. Earlier, he was at the Physics Department of IIT, Chennai for many years. He has worked in low temperature physics and superconductivity.

Absolute zero of temperature is unattainable. One may approach as close to absolute zero as technically possible. In this four part series an attempt will be made to trace developments in the approach to absolute zero. In Part I, I will explain the principles governing liquefaction of gases.

Introduction

Temperature is a measure of the hotness or coldness of an object. It can be measured with a thermometer. The thermometer contains a working substance such as a liquid in a liquid-in-glass thermometer or a wire of some metal in a resistance thermometer. Some property of the substance varies with temperature such as the volume of a given mass of liquid or the resistance of a given piece of wire. To calibrate the thermometer one chooses two standard baths, say a bath of melting ice and a bath of water boiling at one atmosphere pressure. The thermometer is brought into good thermal contact with each bath and the value of the property measured. Arbitrary values are given to the temperatures of the two baths to define a scale of temperature. In the centigrade scale these values are zero for the ice bath and 100 for the boiling water bath. Assuming the value of the property to vary linearly with temperature, one can find the temperature of any other object by bringing the thermometer in contact with the object and measuring the value of the property. It is obvious that the temperature of an object measured using two different thermometers containing different working substances may not agree with each other. Also there is no *a priori* limit for the lowest temperature of an object. The temperature of an object can be negative without any limit. Thus the scale of temperature chosen is not an absolute one.

Temperature scale defined by a thermometer such as a resistance or liquid-in-glass thermometer is not absolute.

All gases under low pressure and high enough temperature obey Boyle's law

$$PV = f(t),$$

where P is the pressure of the gas, V the volume of n moles of the gas and $f(t)$ is a function of temperature. A gas which obeys the above relation at all temperatures and pressures is called an ideal gas. No gas is ideal. Every gas approaches the ideal gas behavior at low pressures and high temperatures. We may define a scale of temperature, called the ideal gas scale, in which we may assume PV to vary linearly with temperature. Experiments indicate that the function $f(t)$ in the ideal gas scale is

$$f(t) = nR(t_c + 273.16),$$

where t_c is the temperature on the centigrade scale and R is the gas constant. If one measures the pressure of a constant volume of gas the temperature values obtained with different gases as working substances will differ very little, especially if the pressure of the gas is low and the temperature is high. To this extent the gas thermometer is superior to the liquid-in- glass or resistance thermometers. However even this thermometer is not absolute as one can choose the zero of temperature at will.

Absolute Temperature Scale

Two remarkable discoveries were made in the study of the efficiency of heat engines initiated by Sadi Carnot, a French engineer. A heat engine absorbs a quantity of heat Q_1 from a high temperature heat reservoir, rejects a quantity of heat Q_2 to a heat reservoir at a low temperature and performs external work W. It is found that heat absorbed at the higher temperature cannot be completely converted into work. Considering an ideal engine in which the processes take place infinitesimally slowly, so that the working substance is in thermodynamic equilibrium throughout the cycle and the cycle is reversible, it can be shown

The absolute temperature scale is defined in terms of the efficiency of a Carnot engine.

that the efficiency, η, of the engine defined as W/Q_1, is independent of the nature of the working substance. So a temperature scale defined in terms of η will not depend on the nature of the working substance in the thermometer. Assuming that a perfect gas is the working substance it is found that

$$\eta = (t_{c1} - t_{c2}) / (t_{c1} + 273.16) ,$$

where t_{c1} and t_{c2} are temperatures of the hot and cold reservoirs on the ideal gas centigrade scale.

One may define a new scale of temperature, the Kelvin or absolute scale, by

$$T = t_c + 273.16 .$$

On this scale the efficiency of the Carnot engine will be

$$\eta = (1 - T_2 / T_1) .$$

The efficiency will approach unity when T_2 approaches zero. The conservation of energy tells us that η cannot be greater than unity. So there is a natural lower limit to the temperature of any object. This is the temperature of the cold reservoir when the efficiency of the Carnot engine will approach unity. On the absolute scale of temperature this will be called the absolute zero.

Unattainability of Absolute Zero

Absolute zero of temperature is unattainable.

Kelvin, who introduced the absolute scale of temperature, realized that the absolute zero is a limiting temperature. One can get very close to absolute zero but one cannot reach it. To reach a low temperature one must remove heat from a system. One can use an ideal refrigerator based on the reverse Carnot cycle. This refrigerator will remove a quantity of heat Q_2 at a low temperature T_2 and reject a larger amount of heat Q_1 at a higher temperature T_1 (say room temperature). This is what any refrigerator does.

Since Q_1 is larger than Q_2, work W must be done on the refrigerant in the process. This is why the domestic refrigerator consumes electrical energy. For an ideal Carnot cycle

$$Q_2/T_2 = Q_1/T_1 = W/(T_1 - T_2).$$

So the work required to remove unit quantity of heat at T_2 is

$$W/Q_2 = (T_1 - T_2) / T_2.$$

As T_2 approaches zero this ratio approaches infinity. It becomes more and more difficult to remove heat as one approaches absolute zero. This is the reason for the unattainability of absolute zero.

The science of production of low temperatures is called *cryogenics*. This is derived from two Greek words, *cryos* meaning cold and *genesis* meaning production. Since the middle of the last century scientists have expended enormous effort in the march towards absolute zero. Cryogenic research was started and sustained by the insatiable urge of man to approach the unattainable and the curiosity of the scientist to find out how matter behaves at low temperatures. This research has resulted in the discovery of several new phenomena occurring at low temperatures. The most well known of these are superconductivity exhibited by some metals, alloys and compounds at low temperatures, and superfluidity exhibited by liquid helium. There are many more interesting phenomena exhibited by solids at low temperatures. As it happens in every scientific endeavour, basic research in low temperatures has led to many applications in such diverse areas as space research, food preservation, medical diagnostics, surgical techniques, high energy accelerators etc. A new and active branch of engineering, called cryogenic engineering, has come into vogue.

The science of production of low temperatures is called cryogenics.

In this series of articles an attempt will be made to trace the development of cryogenics. The aim will be to explain crucial concepts which have contributed to the development of this

branch of physics. Part I of the series will deal with the liquefaction of gases.

Liquefaction of Gases

The first major step that was taken in the march towards absolute zero was the liquefaction of the so called permanent gases. It is well known that a gas like sulphur dioxide can be liquified at room temperature, whereas a gas like nitrogen cannot be liquified at room temperature however high the pressure may be. This earned gases such as nitrogen, oxygen and hydrogen the appellation of permanent gases. The research of Andrews, Amagat and others showed that a gas can be liquified by the application of pressure only if it is cooled below a temperature characteristic of the gas. This temperature is called the *critical temperature*, T_c. For sulfur dioxide the critical temperature is 430.8 K, which is above room temperature. That is why it can be liquified under pressure at room temperature. The critical temperature of nitrogen gas is 126.20 K. It can be liquified under pressure only if it is cooled below the critical temperature. There are no permanent gases. Any gas can be liquified if it is cooled below its critical temperature. The concept of critical temperature is crucial in understanding how to liquify gases. The critical temperatures of the various permanent gases are given in *Table 1*.

A gas must be cooled below its critical temperature before it can be liquified.

The first step in the liquefaction of a gas is to cool it sufficiently so that its temperature is below the critical temperature. One way of doing this is to pass the gas through a bath of a liquid, the boiling point of which is below the critical temperature of the gas. If the critical temperature is well below room temperature it may not be possible to find a suitable liquid bath. One must seek other methods of cooling the gas.

There are two other processes which may be used to cool the gas. The first is adiabatic expansion and the second is Joule-Thomson expansion (J–T expansion) of the gas.

Table 1 Properties of cryogenic liquids.

(Taken from *Cryogenic Process Engineering.* Klaus D Timmer hauss and M Flynn. Plenum Press. New York, 1989).

	LO_2	LAr	LN_2	LH_2	LHe^4
Critical Temp. T_c (K)	154.6	150.7	126.2	32.976	5.201
Normal BP T_B (K)	90.18	87.28	77.35	20.268	4.224
Triple Point T_t (K)	54.35	83.8	63.148	13.803	-
Density (kg/m^3)	1141	1403	808.9	70.78	124.96
Heat of vaporization (kJ/kg)	212.9	161.6	198.3	445.6	20.73
Specific heat (kJ/kgK)	1.7	1.14	2.04	9.78	4.56

If a gas is enclosed in a cylinder under pressure, it can expand to a lower pressure by pushing the piston out and doing work on the piston, *Figure 1(a)*. If the cylinder is thermally insulated so that no heat enters or leaves the system, the expansion is said to be adiabatic. Conservation of energy requires that the gas must lose an amount of internal energy equal to the work done on the piston. The internal energy of the gas is mainly made up of the kinetic energy of its molecules. Adiabatic expansion of a gas should result in a decrease in the average kinetic energy of the molecules of the gas. From the kinetic theory of gases we know that the average kinetic energy of a molecule of a gas is proportional to the absolute temperature of the gas. So adiabatic expansion of a gas always results in cooling. One uses a piston-cylinder or turbine type of an expansion engine to cool the gas.

Adiabatic expansion always cools the gas.

A second method of cooling is by expansion of the gas through an orifice, *(Figure 1b)*. This process is the Joule–Thomson cooling. Here the gas does not perform external work. The total enthalpy, (see *Box 1* for a definition of enthalpy) rather than the entropy, remains constant in this process. Such a process may result in

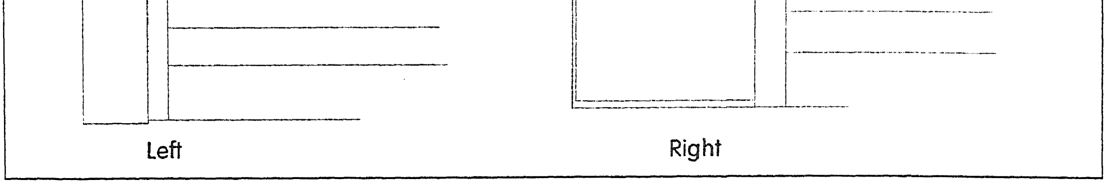
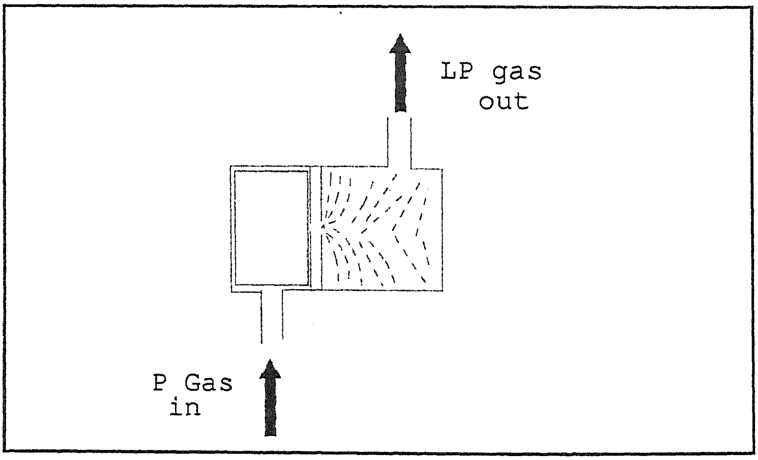


Figure 1 (a) Adiabatic expansion in a piston-cylinder expansion engine. Compressed gas is enclosed between the cylinder and the piston (figure left). The gas pushes the piston out (figure right) with a drop in pressure and increase in volume. If the expansion engine is thermally isolated, this process always produces cooling.

an increase or decrease of temperature depending on the temperature of the high pressure gas. If the temperature of the gas at the inlet of the orifice is more than its inversion temperature, the gas warms up on expansion (see *Box 1* for an explanation). For example helium gas at room temperature undergoing Joule-Thomson expansion at an orifice warms up. Only if the inlet temperature of the gas at the orifice is below its inversion

Figure 1(b) Joule-Thomson expansion. Here high pressure gas expands through an orifice in the plate partitioning the chamber. In this process no external work is done. This process will produce a cooling only if the temperature of the high pressure gas is below the inversion temperature for the gas.



Box 1

The enthalpy of a system is defined as

$$H = U + PV$$

where U is the internal energy and P and V are the pressure and volume of the system. Since the internal energy as well as the work done on the system can be converted to heat, H is called the total heat content of the system. In the Joule–Thomson expansion, no heat enters or leaves the system and no external work is done. So the enthalpy remains a constant in this process. The rate of fall of temperature with decrease in pressure in such a process is given by

$$(\mathrm{d}T/\,\mathrm{d}P)_H = -\{\,(\mathrm{d}U/\,\mathrm{d}P)_T + (\mathrm{d}[PV]\,/\,\mathrm{d}P)_T\}\,/C_p\,.$$

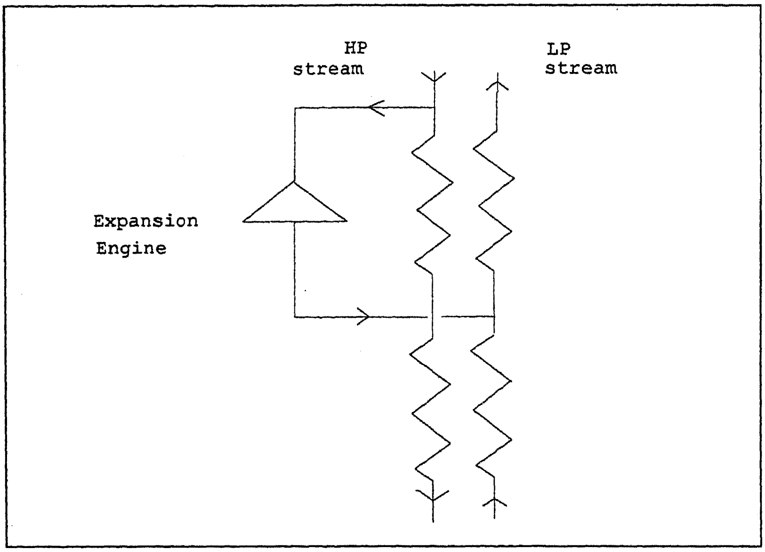
Here C_p is the specific heat at constant pressure of the system. For a gas the internal energy is made up of the kinetic energy of the molecules, which is only a function of temperature, and the potential energy of the molecules due to intermolecular interaction. This interaction is attractive and decreases in magnitude as the molecular separation increases. An attractive energy is negative. As the pressure increases, the molecules come closer to one another and the potential energy becomes more and more negative. So $(\mathrm{d}U/\mathrm{d}P)_T$ is negative. For an ideal gas which obeys Boyle's law PV is only a function of temperature and hence $[\mathrm{d}(PV)\,/\mathrm{d}P]_T$ is zero. An actual gas is less compressible than an ideal gas at high temperatures and more compressible than an ideal gas at low temperatures. This means that at high temperatures $[\mathrm{d}(PV)/\mathrm{d}P]_T$ is positive and at low temperatures it is negative. At some intermediate temperature the value of this derivative cancels the value of $(\mathrm{d}U/\mathrm{d}P)_T$. This is the temperature of inversion. Above this temperature the numerator within flower brackets above is positive. So $(\mathrm{d}T/\mathrm{d}P)_H$ is negative leading to an increase in temperature with a decrease in pressure. The gas will warm up during a J–T expansion if its temperature is above the inversion temperature. On the other hand if the temperature is below the temperature of inversion, the numerator within the flower brackets in above is negative. So $(\mathrm{d}T/\mathrm{d}P)_H$ is positive. The gas cools on isoenthalpic expansion.

temperature will the gas cool on expansion. For helium gas the inversion temperature is below 40 K.

J–T expansion cools the gas only if the temperature of the high pressure gas is below its inversion temperature.

One can use adiabatic expansion or J–T expansion to cool the gas. Such a cooling has to be used in a regenerative way so that the expanded cold gas is used to cool the incoming high

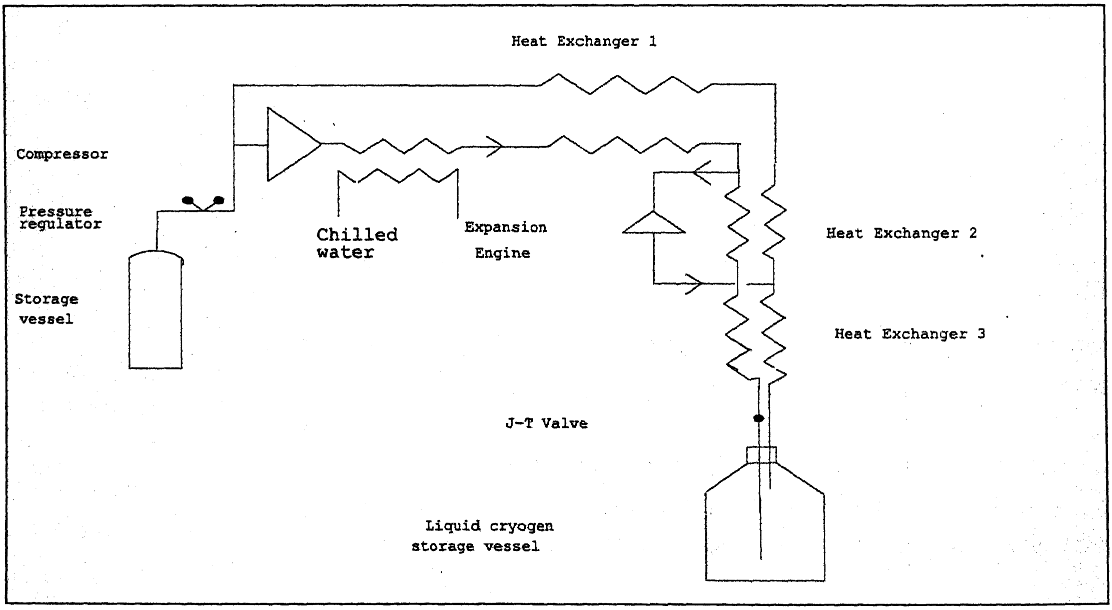
Figure 2 Regenerative cooling produced by an expansion engine in combination with a heat exchanger train. Part of the high pressure gas entering the heat exchanger is diverted to the expander. It expands there to a low pressure and cools. The low pressure gas joins the returning low pressure stream in the heat exchanger. This cold low pressure gas cools the incoming high pressure gas in heat exchanger 1. Thus the temperature of the high pressure gas entering the expander is progressively cooled.



pressure gas. In this way the temperature of the gas after expansion progressively falls till the gas liquifies on expansion.

This regenerative cooling is achieved by the use of a heat exchanger. A counter- flow heat exchanger with an expansion engine is shown in *Figure 2*. A part of the high pressure gas

Figure 3 Schematic diagram of a liquefier. Read text for explanation.



stream is expanded adiabatically. The cold low pressure gas stream flows in one arm of the heat exchanger while the rest of the high pressure stream flows in the other arm. The two counterflowing streams exchange heat. High efficiency heat exchangers are essential parts of a liquefier. *Figure 3* shows a schematic diagram of a liquefier. High purity gas at atmospheric pressure from a storage tank is compressed. The heat of compression is removed by circulating chilled water in the cooler. The high pressure gas enters a train of heat exchangers. Part of the high pressure stream is diverted to pass through an expansion engine. The expanded gas in counterflow cools the rest of the high pressure stream. The final stage in a liquefier is always a J–T expansion stage. On expansion, a part of the expanded gas (about 8 to 10%) liquifies and collects in the storage vessel. The rest of the cold gas returns to the compressor via the heat exchangers. The density in the liquid state is very much higher than the density in the vapour state. So when a part of the gas liquifies the volume of the gas is reduced suddenly. If an expansion engine is used it will result in knocking of the piston on the cylinder. This will cause problems for the mechanical design of the expansion engine. This is the reason why the final expansion stage is a J–T expansion through

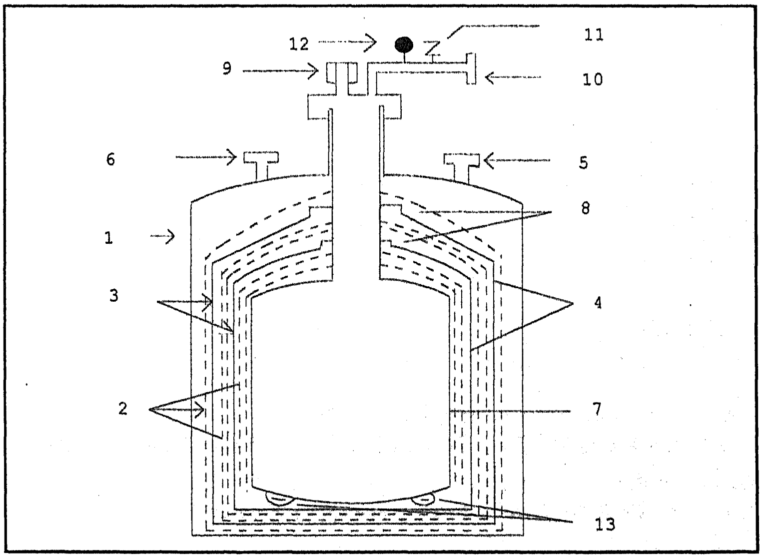


***Figure 4** Storage vessel for liquid cryogen. (1) Outer vacuum vessel, (2) Layers of superinsulation (these are thermally anchored at different points along the neck tube), (3,4) Copper radiation shields anchored at different points along the neck tube, (5) evacuation and seal off port, (6) bursting disc, (7) thin walled SS storage vessel, (8) copper rings soldered to the neck tube to which the radiation shields are anchored, (9) Wilson seal for transfer tube, (10) recovery connection if liquid helium is stored or vent to atmosphere if liquid nitrogen is stored, (11) pressure relief valve, (12) pressure gauge, (13) adsorbent charcoal.*

an orifice. With suitably designed liquefiers air was first liquified by Cailletet in France in 1877, hydrogen by James Dewar in U.K. in 1898, and helium by Kamerlingh–Onnes in Leiden in 1908. It appears that Kamerlingh–Onnes obtained his helium gas from the monazite sands of North Carolina, USA to produce small quantities of liquid helium. Now commercial large scale liquefiers producing hundreds of liters of liquid helium are available.

The normal boiling points of the various cryogenic liquids are given in *Table 1*. Liquid helium has a boiling point of 4.22 K.

Cryogenic liquids have to be stored in evacuated insulated storage vessels.

Cryogenic liquids have to be stored in special vessels so that the heat leaking from the ambient to the cold liquid is a minimum. These are double walled evacuated vessels as shown in *Figure 4*. Evacuation of the space between the inner and outer vessels reduces heat leak by convection. The inner vessel is surrounded by layers of aluminized mylar to reduce heat leak by radiation. The inner vessel in which the liquid is stored is suspended by a thin-walled stainless steel tube to reduce heat leak by conduction. Stainless steel storage vessels of different capacities with about 2% evaporation loss per day are commercially available. With such storage vessels 100 liters of liquid helium can be stored for about 40 to 45 days.

Today closed circuit refrigerators, with limited refrigeration capacity, operating on the Gifford–McMahon cycle, are available to reach low temperatures. With these refrigerators one can perform low temperature measurements down to 10 K in the laboratory.

Address for Correspondence
R Srinivasan
Visiting Professor, Raman Research Institute
C V Raman Avenue
Sadashivnagar
Bangalore 560080, India
rsv@rri.ernet.in

Suggested Reading

- **R B Scott.** ***Cryogenic Engineering.*** **Van Nostrand, 1959.**
- **G K White.** ***Experimental techniques in low temperature physics.*** **OUP, 1959.**
- **A J Croft.** ***Cryogenic laboratory equipment.*** **Plenum, 1970.**

Life Complexity and Diversity

6. Whither Diversity

Madhav Gadgil

The evolutionary history of life is one of continual expansion, of coming into being of increasingly greater diversity of more complex organisms, colonising ever newer environmental regimes. It is a cosmic drama that has become ever more elaborate, as it retires some, but inducts an even larger number of increasingly sophisticated actors into its fold. But the process has by no means been monotonic. It is as if the stage is cleared from time to time to make for fresh beginnings, with major bouts of extinction. Humans are amongst the most complex products of evolution having in turn populated the world with ever growing numbers of complex artefacts. These artefacts are now threatening to overwhelm the diversity of life. But humans may one day engineer life capable of surviving in outer space, and thereby trigger off a new phase of expansion and diversification of living organisms.

Madhav Gadgil who is associated with the Indian Institute of Science and the Jawaharlal Nehru Center for Advanced Scientific Research has not only been fascinated by the evolution of life, but with that of human cultures and artefacts as well.

Mega Extinctions

The most massive episode of extinction of life forms occurred 245 million years ago, at the boundary of the Permian and Triassic periods of geological history. At that time as many as 96% of marine animal species seem to have been wiped out. There have been four other major episodes of extinction in geological history, although none as severe as the one 245 million years ago. The latest of these was 65 million years before the present and it wiped out the dinosaurs (*Figure 1*). These dramatic episodes of mega extinctions were probably due to some geological convulsions; either a hit by an enormous asteroid or a major volcanic eruption.

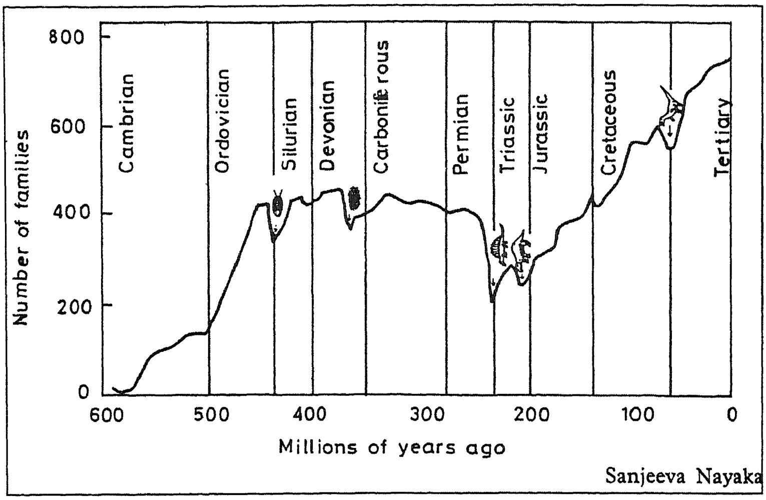


Figure 1 Diversity displays a saw - toothed curve of increase over evolutionary history. While, by and large, more species have been added than deleted from the ecological stage, there have been several catastrophes in which a large fraction of the existing variety of living creatures has been wiped out. Indicated in this figure are the major animal groups that have suffered in such episodes.

Leaving such episodes aside, the background extinction rates have been rather low. Given the relatively complete fossil record, they can be estimated with some accuracy for the last 100 million years. Over this period they have remained at around one species of mammals every 400 years and one species of birds every 200 years. The rates of origins of new species have more than kept up with these, so that the world is now populated by around 9000 species of birds and 4000 species of mammals.

Pinnacle of Complexity

The world is now populated by around 9000 species of birds and 4000 species of mammals.

The warm blooded birds and mammals with high metabolic rates and well developed brains are amongst the most complex of living organisms. This is because complexity resides in diversity of linkages amongst the manifold components of any system. All higher plants and animals are made up of numerous units, the cells. The simplest, like sponges or mushrooms have just a few kinds of cells, more complex organisms like crabs or frogs have tens of different kinds of cells. These cells relate to each other in many different ways; the most advanced of these is through connections of nerve cells. Each nerve cell has many processes which link to other nerve, muscle or gland cells. The brain is of course a bundle of millions of nerve cells, and these can connect to each other in a variety of ways depending on the conditions to which the animal is exposed. This diversity of

possible linkages of nerve cells has conferred on birds and mammals a high level of complexity, leading to a substantial capacity to learn and to adjust their behaivour.

It is the evolutionary trend of development of ever higher levels of complexity that has thrown up our own evolutionary lineage. Humans and their ancestors are notable for the long period of slow growth of the brain, permitting ample opportunity for nerve cells to develop specific interlinkages based on the conditions to which we are exposed. This qualifies humans for the rank of the most complex of living organisms; an animal with an incredibly flexible pattern of behaviour. This flexibility of behaviour, this capacity to learn, has rendered humans tool makers par excellence. Other animals do employ tools; elephants use bits of sticks to scratch themselves, and chimpanzees to draw termites out of the mounds. But the human capacity for production of artefacts far exceeds that of all our ancestors.

Humans and their ancestors are notable for the long period of slow growth of the brain, permitting ample opportunity for nerve cells to develop specific interlinkages based on the conditions to which we are exposed.

Explosion of Artefacts

Members of our ancestral species, *Australopithecus africanus* were already fabricating and using simple stone tools on African savannas 2 million years ago. Their descendants *Homo habilis* and *Homo erectus* elaborated these tool industries further, although at a very slow pace (*Figure 2*). *Homo sapiens* achieved their present day mental capabilities, with a complex symbolic language, some time between 50 to 100 thousand years ago. This permitted them to fabricate increasingly complex artefacts and put them to a variety of uses. Such uses not only include purely practical ones such as sticks for digging tubers, but also those that acquire meaning only in a social context, such as necklaces of shells.

Artefacts, like living organisms are replicating entities. One of them can catalyse the production of many more. As a result, artefacts too have been evolving (*Figure 3*). Artefacts help people improve their access to resources, enhance their social status and dominate others. Bigger and more complex dams, cars, or guns all serve such purposes. So people have been continually fabricating

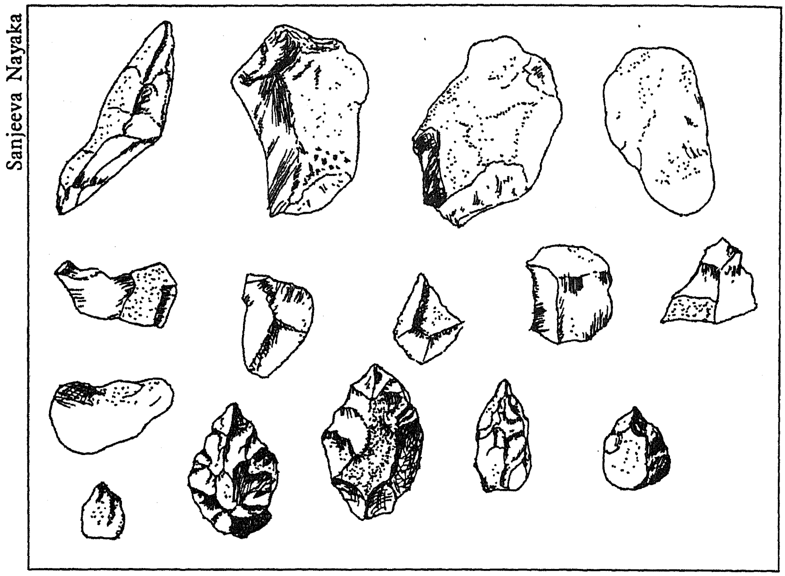


Figure 2 Beginning two million years ago our ancestors began fabricating tools, such as these simple stone scrapers. For hundreds of thousands of years the forms of these artefacts changed but little.

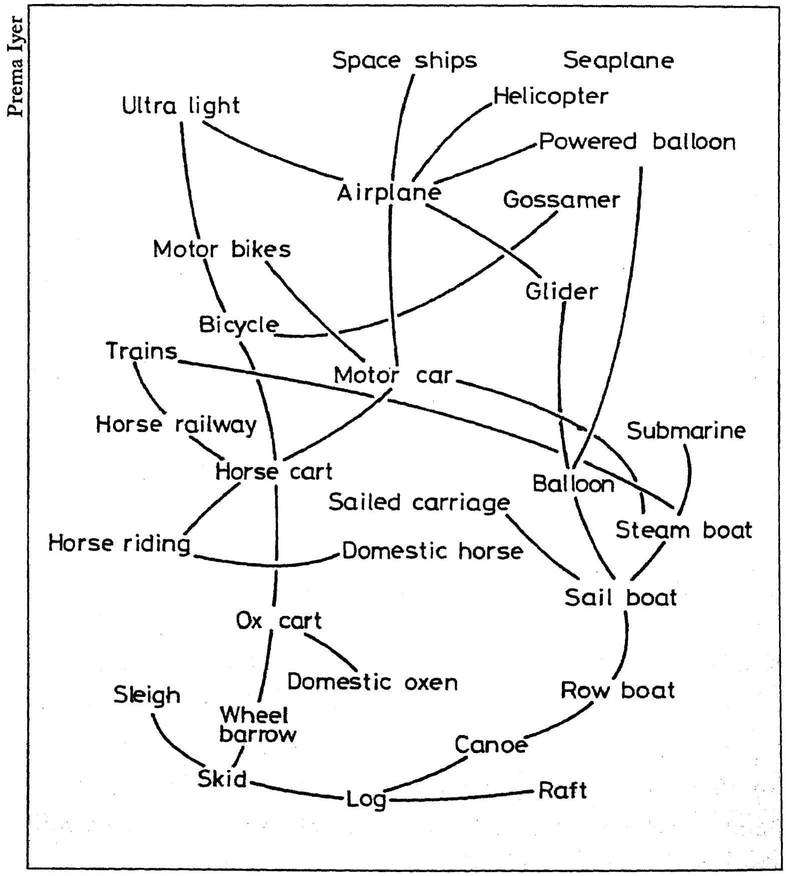


Figure 3 Human artefacts have evolved rapidly over the last fifty thousand years. But unlike living organisms their evolutionary history constitutes a web rather than a bush.

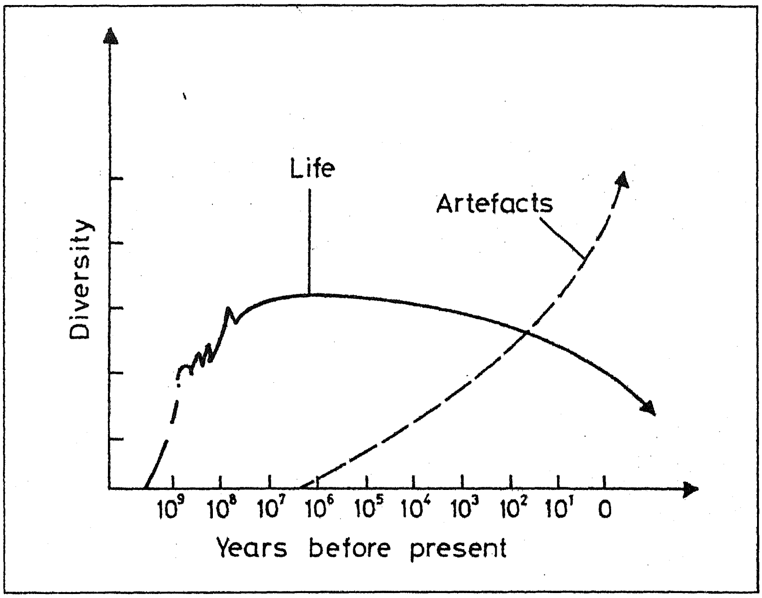


***Figure 4** A schematic depiction of the history and future of diversity of living organisms and of artefacts fabricated by humans, the most complex of the living organisms.*

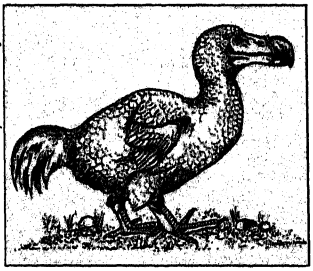
larger and more complex artefacts. Growing populations of these artefacts have been competing with the natural world, eroding the variety of living organisms, even as the diversity of artefacts is continually on the increase (*Figure 4*).

Over the last three centuries, for instance, the rates of extinction of birds and mammals have increased by a factor of fifty compared to the background rates, and it is feared that between 10% to 25% of all living species will go extinct in the next few decades (*Figures 5 and 6*).

Living Artefacts

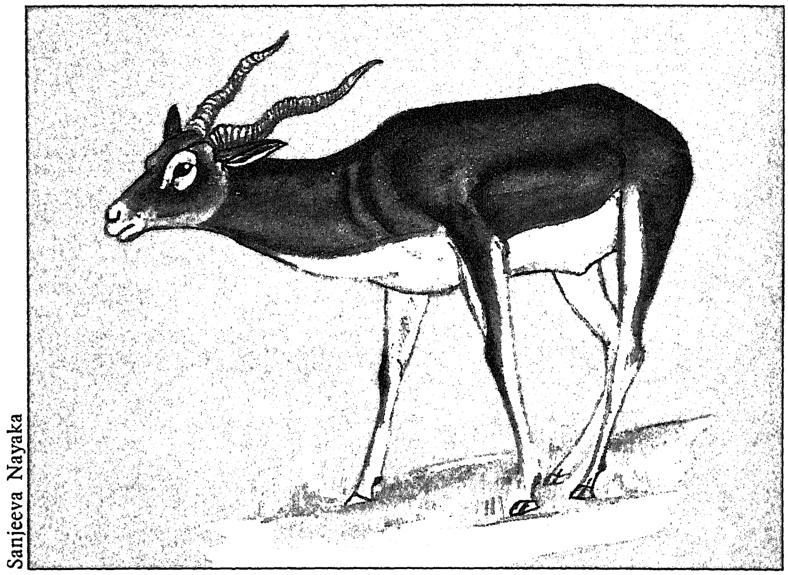
Without doubt then much of human influence on the biosphere has been destructive. But ours is also the only species capable of consciously understanding its impact on nature and doing something about it. As a result, human beings have not only destroyed, but deliberately promoted the diversity of life. The wolf was among the first living species to be moulded by human design. It gave rise to the domestic dog with many new physical and behavioural traits fitting its role as a hunting companion and a

***Figure 5** The Dodo lost its ability to fly in the absence of any large predators on its island home. Human colonizers of oceanic islands have hunted many such flightless birds to extinction over the last several centuries. The Dodo was amongst the last to go; in fact, a painting suggests that Emperor Jehangir possessed a live specimen in his aviary.*



Sanjeeva Nayaka

Figure 6 Blackbuck once roamed the Indian countryside in tens of thousands. Today the species has been driven to the brink of extinction by competing demands on land and by hunting using the gun and the jeep.



guard at camping sites. With time there developed enormous varieties of dogs, widely varying in size, form and behavioural attributes. These may then be thought of as living artefacts (*Figure* 7). Over time humans similarly brought under control and moulded the characteristics of hundreds of other plant and animal species (*Figure 8*). Today these living artefacts cover vast tracts of earth - as paddy fields or teak plantations, as fishponds or herds of cattle and sheep.

Figure 7 Humans have promoted diversity in many species that have been domesticated. The earliest and the most varied of these is the dog.

These living artefacts differ only a little from natural life forms. The last two decades have however witnessed a qualitative increase in the human capacity to manipulate living organisms. Techniques of molecular biology now permit us to move pieces of DNA from one organism to another, creating entirely new forms of

Sanjeeva Nayaka

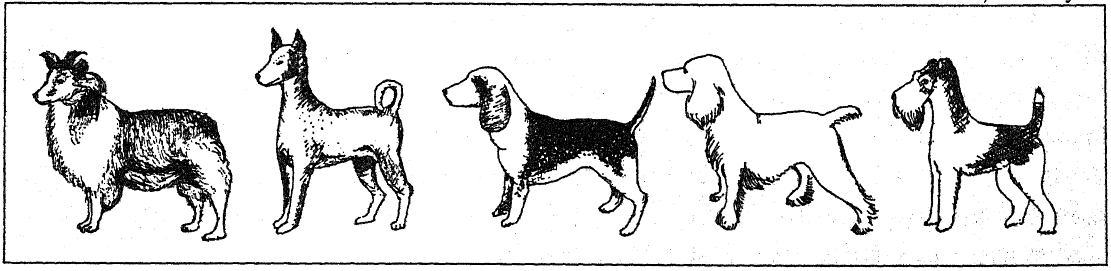
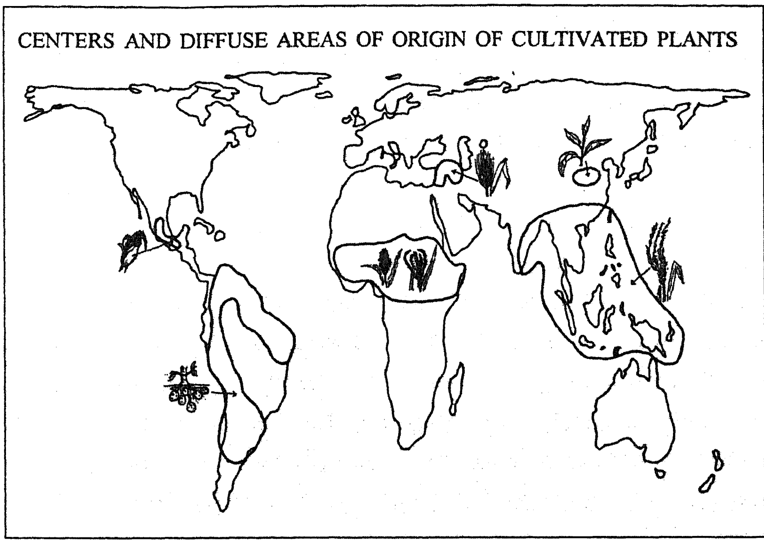


Figure 8 Plants and animals have been brought under domestication by humans in a few major centers in the world; the middle east, and the Mesoamericas being amongst the most important of these. The Indian subcontinent has served as a secondary, yet significant centre of domestication of a number of species of plants and animals. Rice and water buffalo might have been brought under domestication in India, although China and Southeast Asia too have claims as possible centers for origin of these important species. The humped Zebu cattle was domesticated in India, although several centuries after the original domestication of cattle in the middle east. Yak and mithun, two species allied to cattle were domesticated in Western and Eastern Himalayas. North Indian red jungle fowl gave rise to the domestic chicken. Several species of legumes, such as chick pea and green gram, oil seeds like sesame, spices like pepper and cardamom and fruits like mango and jackfruit are other significant contributions of India to the global stock of genetic diversity of husbanded plants and animals.



Sanjeeva Nayaka

transgenic creatures. Thus we can transfer the gene responsible for the production of insulin from a horse to yeast, and that for the production of luciferin, the luminescent substance in glowworms, to the tobacco plant. These developments have enhanced human capabilities for putting other organisms to use - and of making a profit out of them.

The rapid erosion of diversity of life on earth was one of the focal themes at the Earth Summit held in Rio de Janeiro in 1992 and has resulted in the International Convention on Biological Diversity.

Biodiversity Convention

This appreciation of the applied potential of genetic engineering has greatly increased concern over rapid erosion of diversity of life on earth. This was one of the focal themes at the Earth Summit

All nations are expected to organize inventories of the diversity of life within their territories, to monitor what is happening to such diversity, and to develop strategies for conserving it. This is a major scientific challenge, especially for a megadiversity country like ours.

held in Rio de Janeiro in 1992 and has resulted in the International Convention on Biological Diversity. By now 150 countries including all major nations, with the exception of USA, have become parties to this convention. India, China and almost all our Asian neighbours are members of this convention. The convention accepts the sovereign rights of all nations over their genetic resources and promises countries of origin of such resources special privileges when those resources are put to use. At the same time, all nations are expected to organize inventories of the diversity of life within their territories, to monitor what is happening to such diversity, and to develop strategies for conserving it. This is a major scientific challenge, especially for a megadiversity country like ours. We hope that India's community of biologists, old and young, whether they work for, teach or study in surveys, research institutes, universities or colleges, will turn this challenge into a major opportunity.

Future — Bleak or Bright?

But what does the long term future hold for life on earth? There are two alternatives scenarios. The pessimistic scenario, which has a wide following, runs thus: It visualizes a world increasingly polarized between the north and the south; with the north controlling much of the world's wealth as well as technologies, with a stable human population, but exploding material consumption; and the south, harbouring most of the world's poor people with rapidly growing numbers, possessing very low levels of technology. But it is the south that has much of the world's store of biological diversity. This the pessimists expect will be rapidly decimated as the growing populations of the poor degrade the environment at an ever accelerating pace in their attempts to eke out a living. In the meanwhile the northern industrial corporations will drain the biological diversity resources of the south to be maintained in ex-situ collections under their own control to serve as the raw material for future technological developments. Once this has been accomplished the biotechnology industry will in no way be concerned with the destruction of nature in the south. But in their greed these industries would push for the release of

genetically engineered organisms with little forethought. One or other of these organisms would eventually turn against all natural life, of the north as well as the south and wipe out the wonderful diversity of living organisms slowly built up over 3.5 billion years of evolution.

Today's world holds few optimists; but one of the most remarkable of these is the well known physicist, Freeman Dyson. He asks us to contemplate on the history of life on earth, originating in some one obscure place, in the warm shallow seas. Life has gradually expanded from these humble beginnings to occupy all of the oceans, all of the freshwaters, most of the land and the lower reaches of the atmosphere. But today it is confined to the earth, and possibly to a few other planets. It is the open spaces of the universe that are now beckoning life to occupy them. These, life shall occupy with the help of humans, the most complex, albeit the most destructive of organisms that has so far come into being. Freeman Dyson has worked out the attributes of life that can sustain itself in cosmic space. Dyson believes that humans will eventually engineer such organisms and send them out in space. Perhaps they will be prompted to do so by another episode of mega extinction, this time brought about by human interventions, rivalling what happened at the boundary of the Permian and Triassic periods. Out in the interstellar space, the specially engineered organisms will grow, multiply, evolve, spread out and ultimately bring the whole cosmos to life. That would no doubt be a fitting finale to the cosmic drama that began here on this obscure planet three and a half billion years ago.

Life has gradually expanded from these humble beginnings to occupy all of the oceans, all of the freshwaters, most of the land and the lower reaches of the atmosphere.

Suggested Reading

- Dyson, Freeman. *Infinite in All Directions*. Harper and Row. New York. p. 321, 1989.
- Myers, Norman. *The Gaia Atlas of Future Worlds*. Doubleday. New York. p. 190, 1990.
- Gadgil, Madhav. *Genes, memes and artefacts* F di Castri and T Younes (ed.). *Biodiveristy, Science and Development*. CAB International. London. pp. 80 - 90, 1996.

Address for Correspondence
Madhav Gadgil
Centre for Ecological Sciences
Indian Institute of Science
Bangalore 560012, India.
Email:madhav@ces.iisc.ernet.in
Fax: (080) 334 1683

Algorithms

5. Data Types and their Representation in Memory

R K Shyamasundar

R K Shyamasundar is Professor of Computer Science at TIFR, Mumbai and has done extensive research in various foundation areas of computer science.

In this article, we show how the general abstraction of a program as being composed of data and algorithms aids in the understanding of the universality of computers and the development of programs. We describe a simple organization of the memory unit of a computer,and discuss number representation. Subsequently, we show the need and use of *types* in programming. Further, we describe simple data types and illustrate the use of simple data-structures in the development of programs.

Introduction

In the first article of this series, we had compared a *recipe* to an algorithm and the ingredients for the recipe to the input. On similar terms, in the context of programming the role of ingredients is played by the input or the *data*. The algorithm (which is a sequence of instructions) operates on the input and yields an output. We could treat a program as composed of two parts: algorithm and data. Such an abstraction makes it possible to equate the definitions of a computer and a program; in other words,a computer can be treated as consisting of an algorithmic part and a data part. The data part is nothing but the registers and the memory words; these registers or the memory words can store binary values. The algorithm part is the finite set of algorithms built into the computer corresponding to the logic-circuit rules that are used to interpret the stored data as instructions and fetch data. At the most fundamental level the computer works on only one kind of data, i.e., binary bits. The built-in algorithms are fixed and computers can only execute as per these instructions and nothing else.

At the most fundamental level the computer works on only one kind of data, i.e., binary bits.

Programmers rarely write programs in terms of bits. It is more often the case that people would like to use data in terms of the objects they are imagining such as integers, real numbers, character strings, list of items, matrices etc. Further, the algorithms written/contemplated are more varied than the fixed set of algorithms mentioned above. The question that arises is: How can a spectrum of problems be solved by a single machine that executes according to the fixed set of algorithms built into it? The answer is that a computer is a truly general purpose device;its behaviour gets transformed to the behaviour of the program given to it. In other words, the behaviour of the programs are mimicked. The underlying principle lies in the *Stored Program Concept* set forth by John von Neumann. In this model, a stream of information to the computer is interpreted as data at one instant and as an instruction (program) at another instant. It is easy to visualize a simple program that repeatedly takes a given program and its data and executes as per the given

A computer is a truly general purpose device;its behaviour gets transformed to the behaviour of the program given to it.

Stored Program Concept

The stored program concept can be seen in the memo written by J von Neumann proposing a stored-program computer called EDVAC (Electronic Discrete Variable Automatic Computer).

The intention of the proposal was to build a machine that was fully automatic in character, independent of the human intervention once the computation starts. For this purpose, it must be quite evident that the machine must be capable of storing in some manner not only the digital information of the input and intermediate values but also the instructions which govern the actual routine to be performed on the bit stream data. In a special purpose machine these instructions form an integral part of the device and constitute a part of its design structure. However, for a general purpose machine it must be possible to instruct the device to carry out any computation that can be formulated on bit streams (or numerical terms). Hence, the program instructions must be stored in some part of the machine and there should be a part in the machine that can interpret these instructions and execute them. In other words, a general-purpose machine has algorithms that take data and program (as data) and execute the instructions as per the program using the given data and storing the intermediate results in its store and indicating completion after the completion of the program.

The general basis of the universality of computers lies in the notion of Universal Turing Machines proposed by A M Turing.

program on the data given; as soon as the program terminates, it is ready to interpret the next data (which is again a program and its data). The general basis of the universality of computers lies in the notion of *Universal Turing Machines* proposed by A M Turing and will be discussed in forthcoming articles.

In general, a program is composed of objects familiar and convenient to the user. This process enables the design and construction of complex programs. A programmer arrives at a solution starting from a skeletal solution (which may not be suitable for execution on a computer) to an implementation. In other words, a programmer goes through a series of abstraction layers before arriving at the final program. In each layer, the programmer attends to the details of some part ignoring the presence of others. This process of abstraction is essential for the construction and understanding of the programs. The main purpose of abstraction is to concentrate on similarities and de-emphasize the differences. By doing so, one would be concentrating on the relevant features of the problem at hand, ignoring the irrelevant features of the problem. For instance, when trying to arrive at a payroll program the financial aspects of an employee are more important rather than his/her physical or metabolic features. In a similar way, while analyzing the engine-performance of an automobile, its colour and aesthetic features are irrelevant. It can also be seen that abstractions need not be unique. Depending on the representation/manipulation of data one could arrive at different abstractions. Abstraction is the key to gaining intellectual mastery over any complex system. Needless to say it requires great skill and experience to use abstraction effectively. Abstract notations at higher levels require translators that take source programs (in high-level languages/notations the users would have written)and translate them into programs in terms of bits that can be executed by the computer. Program abstractions can be broadly divided into control structure abstractions, procedural abstractions and data abstractions.In earlier articles of this series,we have had a look at control

structure abstractions and procedural abstractions. In this article, we will consider data abstractions.

A variable is not a data item; it is rather a name.

Types

In the previous sections, we have used inputs, outputs and other intermediate computation objects, in a rather intuitive way without worrying about their structure. For instance, one could think of the procedure for summing N numbers as corresponding to summing the salaries of N employees. These items (numbers in this example) are generally referred to by the generic term *data*. Some of the questions that we need to look at are: How do we refer to the data? Is there a need for *type* information along with the data? How do we structure data? These aspects are discussed in the sequel.

Constants and Variables

In programming languages, facility of referring to an item through a name is very common. Some of these named objects will have the same value throughout the whole program execution. These items are referred to as *constants*. Some named objects take on new values as the computation progresses. These items are referred to as *variables*. For instance, in the example of 'sum-of-N-numbers', we used *sum* to accumulate the partial and final results. It must be emphasized that a *variable* is not a data item; it is rather a *name*. In short, the name of a constant or a variable is only a mnemonic aid to the programmer; it has no meaning to the computer. The identifiers are like names to locations — such as a number given to a post box in a post-office. However, there is one difference: there can possibly be more than one name for the same location (similar to *nicknames*). That is, *aliasing* is a possibility; there cannot be a hotel room with two room numbers but two *names* can refer to the same location. The translator of the program that maps the text into a bit stream merely associates with each identifier a unique *location* in the memory of the

computer;the *location* is usually referred to as the *address* of the variable. Thus, when an instruction uses a constant or a variable named through an identifier say *sum*, the computer fetches whatever value is in the location corresponding to *sum*. Various operations can be performed on *expressions* formed from constants and variables. One could say " set x to $x+1$ ", " set x to $x+y$" and so on. One can also test whether "$x=100$ " or "$x>100$ " holds.In other words, naming constants and variables are similar to the use of symbolic variables and expressions in algebra.However, one has to note an important difference. In the algebraic expression "$a+b$ ", it is implicitly understood that we can add numbers denoted by a and b. However, in the context of computers where *unambiguous representation is a must*, we need to use additional information (i.e., *type*) of the variables or constants used. For example, we can talk of the predecessor of a natural number but not of a real number. Similarly, if one is keeping track of the name of an employee, we know that there is no need to do any arithmetic operations on the names of the employees. In traditional algebra, we understand the type of the variable from the context implicitly. In the context of computers, it is essential to keep track of the type of the data *explicitly* so that the data can be interpreted unambiguously. This aspect will become clear if we look at the basic organization of memory and the internal representation of basic types such as integers,reals, and characters.

In the context of computers, it is essential to keep track of the type of the data explicitly so that the data can be interpreted unambiguously.

Memory Organization and Internal Representation of Data

Let us consider a simple organization of the main memory unit of a computer. A simple organization of computer memory has the following features:

- The memory unit is subdivided into sub units each of which can be individually accessed. The subunits are grouped into larger units each of which can store the same amount of information (some fixed number of bits). The organization is

referred to as digit-or character-organized depending upon whether such a unit has the capacity to hold a digit or a character respectively. In the current day architectures,such an unit is usually referred to as *byte* and consists of eight bits. For the purpose of carrying out arithmetic or logical operations the memory is organized in terms of a fixed number of bytes (2, 4 etc.) referred to as *words*. Such an organization is usually referred to as word organized.

- Each of these subunits/bytes/words has an associated number called the *address* (or location) that can be used for accessing or locating it.

Let us look at the internal representation of some of the basic data items assuming the memory of a computer to be word organized where each word is comprised of four bytes and each byte consists of 8 bits. We will confine ourselves to this simple organization while discussing the internal representation unless otherwise stated. In this context, let us see how common data types such as numbers, characters and real numbers are represented.

Memory is organized in terms of a fixed number of bytes (2, 4 etc.) referred to as words.

Representation of Numbers

Everyone is familiar with radix notations; in practice we use radix 10 and in the binary system of representation the radix used is 2.For example, an integer 75 (without any sign) can be represented in 32 bits by (the space between bytes has been used only for ease of reading):

$$75 \equiv 00000000\ 00000000\ 00000000\ 01001011$$

For representing integers, we should have the ability to represent positive as well as negative numbers. There are various methods of representing negative numbers; the choice influences the way the arithmetic operations are carried out. The simplest one is to represent the magnitude and the sign explicitly. For example,

$$-75 \equiv -00000000\ 00000000\ 00000000\ 01001011$$

The major difference between signed magnitude and 2's complement notation in practice is that shifting right does not divide the magnitude by 2.

This is called the *signed-magnitude* representation. Naturally, the representation is appealing as it coincides with the classical notational conventions. A potential disadvantage is that minus zero and plus zero can both be represented while they should usually mean the same number; this requires care particularly when one is computing through mechanical or automatic machines.

One of the commonly used notations for representing negative numbers on the electronic computers is the *two's complement notation*. For instance, in the context of 32 bits as above, if we subtract 1 from 00000000 00000000 00000000 00000000 we get 11111111 11111111 11111111 11111111 which denotes –1 without an additional sign and computations are done modulo 2^{32}. Thus, –75 in this notation would be represented as

$$-75 \equiv 11111111\ \ 11111111\ \ 11111111\ \ 10110101$$

The major difference between signed magnitude and 2's complement notation in practice is that shifting right does not divide the magnitude by 2. A disadvantage of 2's complement notation is that it is not symmetric about zero; the largest negative number representable in p digits is not the negative of any p – digit positive number. Another notation that is used is the one's complement notation. The one's complement of a number is the result of changing each zero to one and one to zero. The two's complement can be defined as the one's complement plus a one in the least significant position. In fact, the two's complement and one's complement correspond to the ten's complement and the nine's complement used in decimal notation respectively.Computers use these notations for reasons connected with circuit economy and efficiency.

Real Numbers

Real numbers are approximated as decimal fractions. Numbers are represented in two forms: fixed-point and floating-point representations. A fixed-point number in a computer is one for

which the computer always assumes the binary or the decimal point to be at the same place for numbers. This place is either at the left-hand end of the number (all numbers considered to be less than 1 in magnitude) or at the right-hand end (all numbers are treated as integers). For understanding the problems of this representation, consider the addition of two decimal numbers, -65.31 and 78.42. Now, if we assume that the decimal points are assumed at the right-hand end, then adding the two numbers, we get 1311;the decimal point can be placed without any problem and we get the result 13.11 as anticipated. Now consider the addition of -65.31 and 784.2 . Again,assuming the decimal point to be at the right-hand end,we get the same answer (i.e., 1311) as in the previous case which is clearly wrong. To overcome this problem, it can be easily seen that one number has to be shifted relative to the other so that the decimal points are lined up one above the other (which we do in the paper-pencil method);this process is referred to as *scaling*. Scaling is somewhat tedious in lengthy computations. Another problem that arises in this representation is the difficulty in handling numbers that differ very widely. For instance,in a 36-bit word machine (sign and 35 bits), if one wants to add 10^8 and 10^{-4}, it is not possible to do it in fixed-point arithmetic because no-matter where the imaginary scaled binary point is, the two numbers can never differ by more than 2^{35}. As a result of these scaling and magnitude problems, a new representation called the *floating-point* representation was developed.

A floating-point number is a sequence of bits which is interpreted to have two distinct parts, one called the exponent, *E* and the other called the fractional part (often referred to as the mantissa), *M*.

A floating-point number is a sequence of bits which is interpreted to have two distinct parts, one called the *exponent*, E and the other called the fractional part (often referred to as the mantissa), M. The number is interpreted as $M \times 2^E$ in the binary internal representation; in the decimal representation the number is interpreted as $M \times 10^E$. For example, a 10-digit representation of 86.3 is

02 86300000

which is interpreted to have $E = 2$ and $M = .863$, i.e., 0.863

$\times 10^2$ where it is assumed that the decimal point for the fractional part is at the left-end of the mantissa. It can be easily seen that $E = 3$ and $M = .0863$ is also an equally correct representation. For this reason, it is required that the first digit (or the bit) to the right of the binary or the decimal point is nonzero unless the mantissa is itself zero. Since we have two parts of the numbers, we should have the ability to place signs to both the parts. In the floating-point representation one bit is used for representing the sign of the mantissa while for representing the sign of the exponent a different technique is used. Assuming 8-bits for exponent, the range of numbers that can be represented in 8 bits is from 0 to $2^8 - 1$. Thus,arbitrarily one assigns 10000000 to correspond to the zero exponent. Then the range of exponents is from -128(denoted by 00000000) to $+127$ (denoted by 11111111). This notation is referred to as the *excess* -128 notation. For instance, an exponent part of 10000110 denotes an exponent of 6 (i.e., 134 -128).

Characters and Digits:

It easily follows that by a proper encoding (or mapping) one can store digits/characters if the memory devices used are all binary devices. Characters are generally given seven- or eight-bit values specified by the American Standard Code for Information Exchange (ASCII) code. For instance,some of the typical codes for some digits and characters are shown below:

ASCII Code	Digit/Character	Base10 Representation
00100000	<space>	32
00110011	3	51
00111000	8	56
01000101	E	69
01010011	S	83
01010100	T	84

IEEE Floating Point Standard

We need numbers of various types such as integers, rationals, reals, irrationals etc. There is a variety of ways of representing nonintegers. One such method is the fixed point method described in the article. In fixed point representation,one essentially uses integer arithmetic operators assuming the binary point to be at some point other than the rightmost point. Though there are several possible representations, floating point representation is widely used. One of the difficulties with floating point representation is that the meanings of the operations are not as simple as that of integer operations. Further, the operational understanding and implementation depends on the number of bits for the mantissa and exponent (e.g., typical notions such as round-off, overflow, underflow etc.).

The main purpose of standardization is to aid *software designers* to develop efficient and reliable software, *computer designers* to develop techniques for implementing efficient computation of the operations on hardware and *computer manufacturers* to arrive at accelerators for floating point evaluations. Towards providing a unified account of the floating point representation, a standard format has been specified by IEEE standard 754-1985. The standard has gained wide acceptability from software designers to computer manufacturers. The merit of using a variant of the standard floating point representation is similar in spirit to that of using floating point representation over other methods of representation.

An important feature of the above standard is that *computation* continues despite *exceptional* situations like dividing by zero or taking a square root of a negative number etc. For instance, the result of taking the square root of a negative number is equated with a special sentinel value N a N (Not a Number) corresponding with a bit pattern for which there is no valid proper number in the domain. Such a representation has the advantage of portability (i.e., the same program can be used over different systems) since at the program level one can detect as to when the result goes outside the range of the computation. The actual action to be taken in such a situation, called "exceptions"(such as divide by zero etc.), is anyway language and operating-system dependent.

The round-off specification is another feature which when developed as per the standard, leads to unambiguous hardware designs. For instance, consider multiplying two numbers 22.1×0.5 (equal to 11.05) which is needed to be rounded off to two digits. The question is: Should it be rounded to 11.0 or to 11.1. As per the standard, such halfway cases are rounded to 11.0 and not 11.1. The standard has actually four *rounding modes*; the default mode is round to nearest and round to an even number in case of ties.The other modes are: round toward 0, round toward $+\infty$ and round toward $-\infty$. Details or references to the standard specifications can be found in the references given at the end.

Need for Data Types

From the above representation of numbers and characters, it should be clear that unless the type of the memory word is known, the data cannot be interpreted properly;in fact, from the table shown it can be easily seen that it leads to misinterpretation. Thus, there is a need for explicit *type* information while writing programs.In general, the type information becomes useful for

- Interpreting the data (and also the instruction).
- Deciding about the memory requirements

Indian Script Code for Information Interchange (ISCII)

There are 17 officially recognized languages in India: Hindi, Marathi, Sanskrit,Punjabi, Gujarati, Oriya, Bengali, Assamese, Telugu, Kannada, Malayalam, Tamil,Konkani, Manipuri,Urdu,Sindhi and Kashmiri. Each language is written in its own script. With the exception of languages such as Urdu, Sindhi and Kashmiri all other languages are written in the following scripts: Devanagari, Punjabi, Gujarati, Oriya, Bengali, Assamese, Telugu, Kannada, Malayalam and Tamil. The above scripts have evolved from the ancient Brahmi script. The official language of India, Hindi is written in the Devanagari script; Marathi and Sanskrit are also written in the Devanagari script; Devanagari script is also the official script of Nepal. Though Urdu, Sindhi and Kashmiri get written in Devanagari (Sindhi gets written in Gujarati script as well), these are primarily written in Perso-Arabic scripts.

ISCII has been adopted by the Bureau of Indian Standards and is intended for use in all computer communication media that requires the usage of 7 or 8-bit character. In an 8-bit environment, the lower 128 characters are the same as the ASCII character set. The top 128 characters cater to all the 10 Indian scripts based on the ancient Brahmi script. A different standard is envisaged for the Perso-Arabic scripts. An optimal keyboard overlay for all the scripts based on Brahmi script has been made possible by the phonetic nature of the alphabet. Such a standard has the advantage of having a common code and keyboard for all the Indian scripts and hence,transliteration between different Indian scripts becomes straightforward. Further details can be found from the document IS 13194: 1991 from the Bureau of Indian Standards.

Some of the basic data types that are commonly used in high level-languages are given below :

- Integers: The internal representation is as discussed above.

- Boolean: The domain of Boolean has two values usually denoted *true* and *false*. The internal representation of this type requires just one bit (1 denoting *true* and 0 denoting *false*). In a word of four bytes, one can also use the representation of all 0's and all 1's for representing *true* and *false* respectively.

- Reals: The internal representation is as discussed above.

- Characters: The internal representation is as discussed above.

- Strings: Strings are nothing but a sequence of characters. Thus, one way of representation is to use a length designator that indicates the length of the string followed by that many characters. For instance, the string "SET <blank space> 83"is represented by 00000110 01010011 01000101 01010100 00100000 00111000 00110011 where the first 8 bits are the length qualifier; in this case the length qualifier has value 6 which is indeed the length of the string (note that <blank space> denotes one character).

- Sets: Consider a 4 byte word each consisting of 8 bits. We can use a word to represent a set consisting of 32 elements; we could map the elements to 0 to 31 positions in the bit representation say, from the least significant to the most significant position. For instance, a set {31,8,4,2,1} is represented by10000000 00000000 00000001 00010110 where a 1 in the ith position indicates that element *i* is in the set; if it is 0 it indicates that the element is not in the set. That is, a map between the element and the position of the bit in the

word is to be maintained. Under such a representation, several operations on sets can be performed efficiently using logical operations.

In addition to the reasons for the type information cited above, the type information is generally useful to associate entities used in the program with their types, for the following reasons:

- It provides clarity to the program by specifying the potential collection of values that may be assumed by the entities.
- It delineates the set of meaningful operations that can be performed on entities of such type.
- It enables the programmer to arrive at a proper representation and manipulation of the entities.

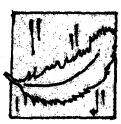
It assists the programmer to prevent or detect errors in the program well in advance during the program design process.

In the next article, we illustrate the use of data structures in designing algorithms for sorting and searching.

Address for Correspondence
R K Shyamasundar
Computer Science Group
Tata Institute of Fundamental Research
Homi Bhabha Road
Mumbai 400 005, India
email:shyam@tcs.tifr.res.in
Fax: 022-215 2110

Suggested Reading

- R G Dromey. *How to Solve it by Computer*. Prentice Hall International, 1982.
- V Rajaraman. *Fundamentals of Computers*. 2nd Edition. Prentice Hall of India Pvt. Ltd, 1996.



No science is immune to the infection of politics and the corruption of power

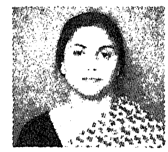
Jacob Bronowski

Courtship in Frogs

Role of Acoustic Communication in Amphibian Courtship Behaviour

Debjani Roy

Vertebrate vocalization came into existence for the first time in frogs. Acoustic signals produced by the frogs have well-defined physical characteristics and a clear biological meaning. The signals are meant to attract and assess the sex, species identity and genetic quality of potential mates. Acoustic communication plays a central role in the courtship behaviour of frogs.



Debjani Roy is at the Institute of Self Organising Systems and Biophysics, North Eastern Hill University, Shillong. She is fascinated with the world of frogs. She has been trying to understand the communication system – mainly their acoustic communication and acoustically mediated behaviour while studying the frogs from north east India.

The chordates originated in the sea as jawless fishes. The bony fishes, which evolved from them are the most plentiful vertebrates today. The first vertebrate land dwellers were the amphibians, but they are not truly terrestrial because they still require frequent access to water. The first true terrestrial vertebrates were the reptiles, which independently gave rise to the birds and to the mammals, including humans (*Figure 1*).

The amphibians – descendants of the crossopterygian fishes were the first land dwellers. Their transition from fresh water to land was a momentous step in vertebrate evolution. The word amphibia is derived from the Greek word *amphibious* meaning *double life* because of their two phase life style: a free living larval aquatic stage and a terrestrial juvenile and adult stage. Contemporary amphibians include caecilians (Order *Gymnophiona*), salamanders (Order *Urodela*) and frogs and toads (Order *Anura*).

Amphibians were the first vertebrates to have evolved a partially terrestrial way of life. This became possible due to a series of anatomical and physiological adaptations to the new environment. One feature of this new environment was the acoustic world – the sounds that were abiotically caused. The amphibians that

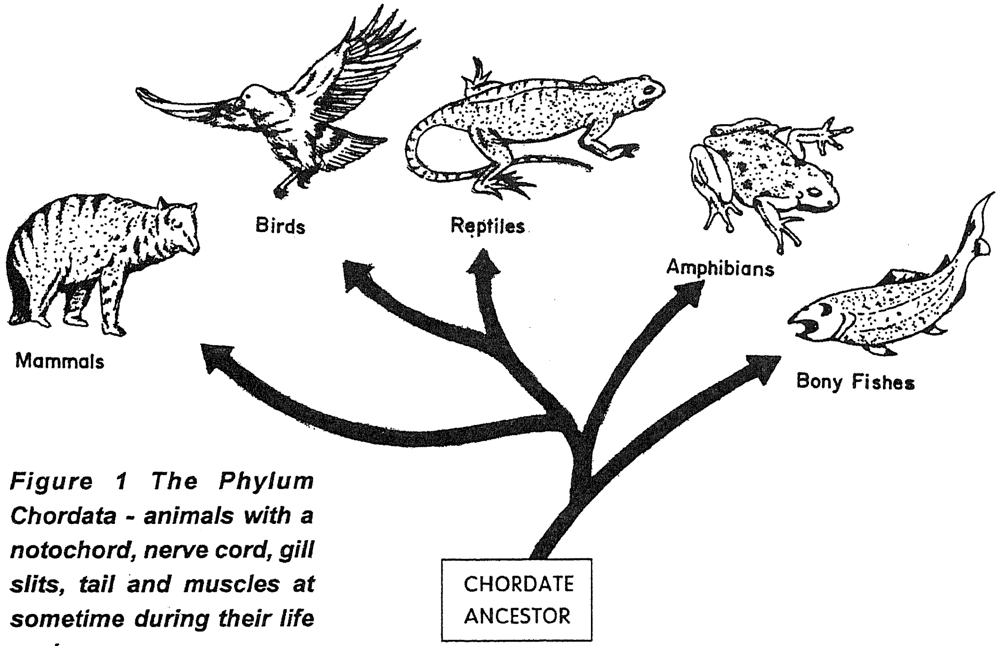


Figure 1 The Phylum Chordata - animals with a notochord, nerve cord, gill slits, tail and muscles at sometime during their life cycle.

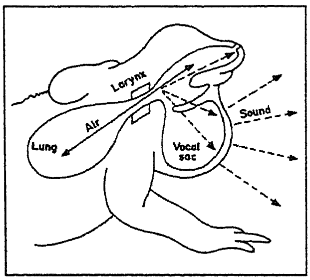


Figure 5a Sound production in frogs. Sound production uses the respiratory ventilation cycle without releasing air to the outside. Before calling the, bucopharyngeal force-pump inflates the lungs and vocal sacs. With the nostrils closed, the body muscles contract, pushing a pulse of air through the larynx, virbating the vocal cords. Sound radiates outwards and the vocal sacs resonate it.

adapted to the changed acoustic conditions of their environment most successfully were the frogs and toads (hereafter referred to collectively as frogs). They adapted so well that acoustic communication came to play a central role in their reproduction. Vertebrate vocalization came into existence for the first time in frogs.

The sound production apparatus of the frogs consists of the larynx and its vocal cords. The laryngeal apparatus is well developed in the males, who also possess a vocal sac. Air from the lungs is forced over the vocal cords and cartilages of the larynx, causing them to vibrate and regulate the frequency of sound. Muscles control the tension of the vocal cords. Vocal sacs act as resonating structures and increase the volume of the sound (*Figure 5a & b*).

The emitted sound consists of regularly alternating compressions and rarefactions of the air, basically increased and decreased air

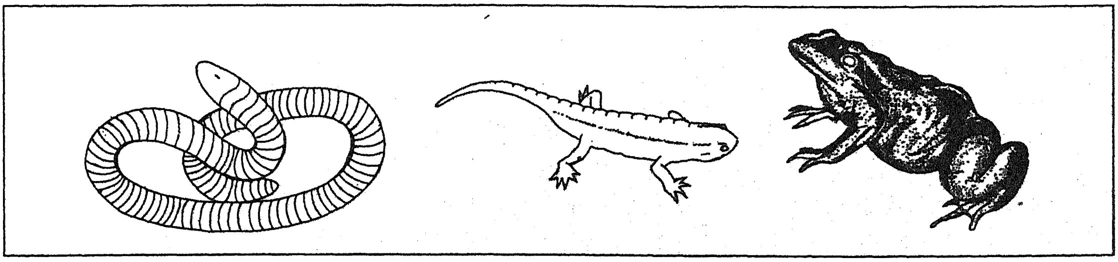


Figure 2 (top left) Gymnophis – Order Gymnophiona.
Figure 3 (top center) Hydromantoides – Order Urodela.
Figure 4 (top right) Rana Species – Order Anura.

pressure respectively forming a simple harmonic or sinusoidal motion. The resulting displacement of the air molecules form a sine wave. Like all sine waves it can be described fully by stating its *amplitude* or height and its *period*, which is the time required to go through one complete cycle. The period is inversely related to the *frequency*, which is the number of cycles per unit time.

Thus,

Period (Seconds) = 1/Frequency (Seconds^{-1}).

Sounds are usually designated by their frequency in cycles per second or hertz (Hz). A thousand hertz is called a kilohertz (kHz) (*Figure 6*).

The *amplitude* of the wave is correlated with its perceived loudness, and a special scale, the Decibel Scale, is used to measure the amplitude of pressure waves.

Figure 5b A calling toad – Bufo melanostictus with its vocal sac inflated.

Sound Pressure Level (SPL) in decibels = 20 $\log_{10} Pt/Pr$ where, Pt is the test pressure and Pr is the reference pressure (2×10^{-4} dynes/cm^2). The background noise in a quiet room is generally about 30 dB SPL; a typical human voice heard at close range is about 60–80 dB SPL; and rock music as commonly played reaches 120–150 dB or higher. Some species of frogs can produce ear splitting sounds of 114 to 120 dB.

Box 1 Amphibian Orders

Order *Gymnophiona*: Limbless, pelvic and pectoral girdles absent; body elongated and cylindrical, regularly encircled by grooves forming segments; head blunt and cone shaped with an overhung lower jaw and nearly invisible, degenerate eyes; tail if present short and blunt. They are usually in shades of pinkish brown and grey. They are fossorial, living by burrowing in soil, a few are aquatic (*Figure 2*).

Order *Urodela*: Lizard like, having low slung bodies with moderate length limbs and long tail; head broad with distinct eyes, separated from the body by neck. The aquatic forms have reduced limbs and appear eel like. They are coloured animals, sometimes with bold patterns. They are found in both cool mountaineous regions as well as hot lowlands, inhabiting slow moving or still water (*Figure 3*).

Order *Anura*: Body dominated by long powerful hindlimbs, an adaptation for their jumping locomotion. The entire body is adapted for such adaptation, with shortened body with broad head, no neck, no tail, with well developed forelimbs propelling synchronously with the hindlimbs. They have a wide range of colourations. A few species being entirely aquatic, most are semiterrestrial to terrestrial and arboreal (*Figure 4*).

For precise evaluation of frog calls for the message and meaning they emit, the sounds are recorded with the help of a professional walkman and unidirectional condenser mircophone and stored on audiocassettes. The sounds stored as waveforms on audiocassettes are next digitized by an analogue to digital converter to get oscillograms, sonagrams and mean spectra by using computerised Fast Fourier Transformations (FFT). The

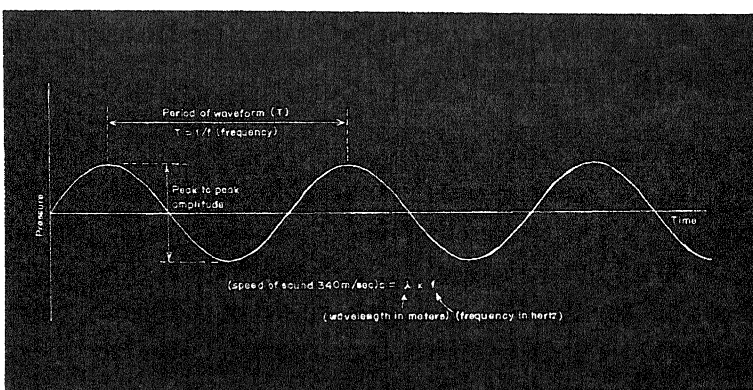


***Figure 6** A sinusoidal wave or sine wave propagates through space, the ambient pressure in the air is measured with a microphone probe at a fixed point. The speed of sound is constant in air (approximately 340m/sec) and is related to both the wavelength (l) and frequency (f) of the wave shown in the equation in the figure. The tympanic membrane of the ear moves in response to alternating condensation (peaks) and rarefaction (troughs) of the sound wave.*

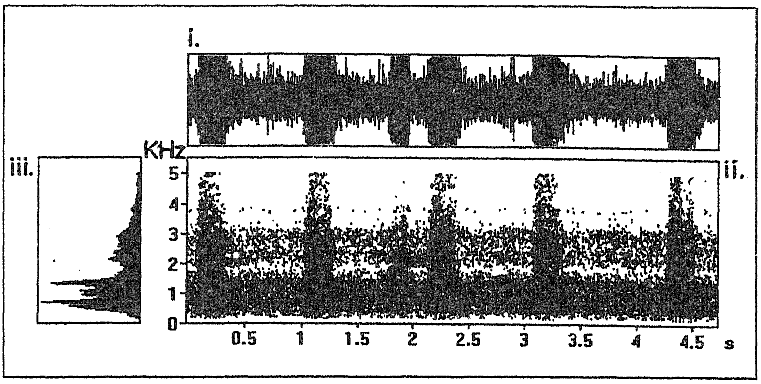


Figure 7a ***Bull frog – Rana tigerina.***

Figure 7b (i) Oscillogram or waveform representation of the advertisement call of bull frog with amplitude in the abscissa and time in the ordinate. (ii) & (iii) Corresponding sonagram and mean spectrum with frequency in the abscissa and time trace in the ordinate with amplitude in shades of grey. The mean spectrum showing frequency spread from 0.13 to 4.91 kHz. The frequency distribution is bimodal with dominant frequency at about 0.69 and 1.34 kHz.

FFT allows one to obtain a frequency-domain representation of a time domain wave form. Its popularity derives from the availability of fast algorithms of its calculations and also from the ubiquity of its use in the analysis of acoustic signals in animal communications (*Figure 7a & bi, ii,iii*).

Most calls last about 2 secs. which may be repeated rapidly in sets. Calls generally lie between 200–500 Hz which is a good broadcast spectrum for these animals, since they can avoid interference from other sounds. Most frog calls span a broad frequency spectrum of 1–2 kHz, whose energy is concentrated in narrow frequency bands either with a single dominant fundamental frequency or first harmonic of the sound or in harmonics. Twice the fundamental frequency is called second harmonic, three times the third harmonics and so on. The sounds produced are not mere sounds or croaks of frogs but have definite physical, physiological and biological meaning. The sound production is a primary reproductive function of the frogs. Male frogs produce species specific complex sounds composed of numerous, closely spaced, often harmonically related components with predominant amplitudes that fall into distinct frequency bands, the location of

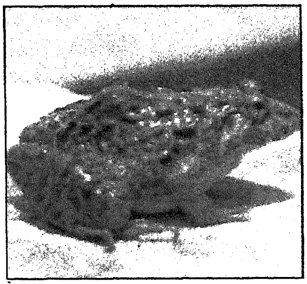


Figure 8a Cricket frog – Rana limnocharis.

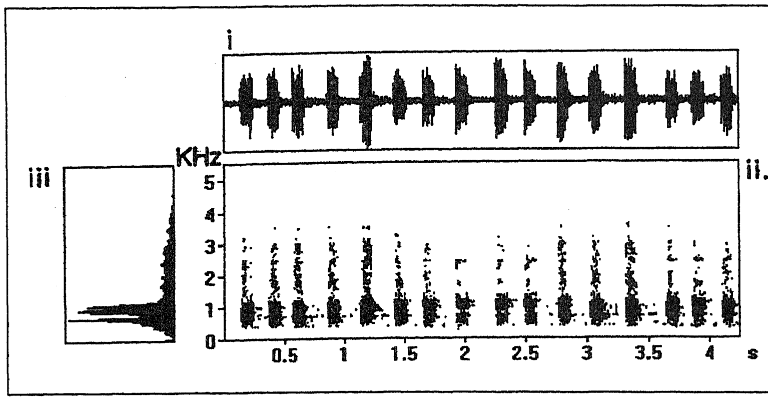


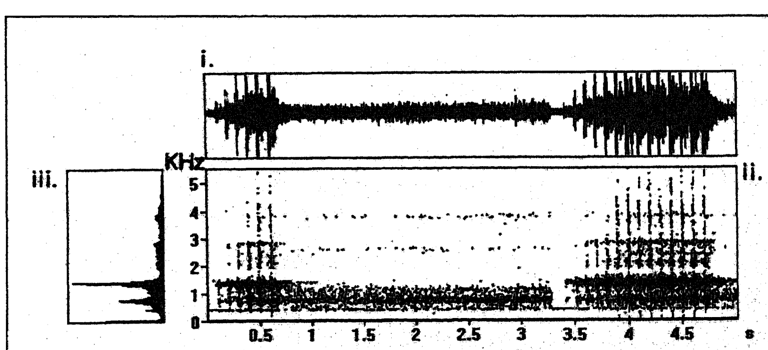
Figure 8b (i) Oscillogram or waveform representation of the advertisement call of cricket frog with amplitude in the abscissa and time in the ordinate. (ii) & (iii) Corresponding sonagram and mean spectrum with frequency in the abscissa and time trace in the ordinate with amplitude in shades of grey. The mean spectrum showing frequency spread from 0.35 to 4.40 kHz. The dominant frequency at 1.08 kHz.

which are characteristic for each species (*Figure 8a & bi, ii, iii; 9a & bi, ii, iii*).

The calls have species specific temporal and spectral characteristics. Notable features of the temporal structure are call duration (sec.), intercall interval (sec.), pulse number, interpulse interval (sec.), call intensity (dB); the spectral part consists of frequency

Figure 9b (i) Oscillogram or waveform representation of the advertisement call of skipper frog with amplitude in the abscissa and time in the ordinate. (ii) & (iii) Corresponding sonagram and mean spectrum with frequency in the abscissa and time trace in the ordinate with amplitude in shades of grey. Mean spectrum showing the frequency spread from 0.33 to 4.36 kHz. The frequency distribution is bimodal with dominant frequency at 0.78 and 1.42 kHz.

Figure 9a Skipper frog – Rana cyanophlyctis.



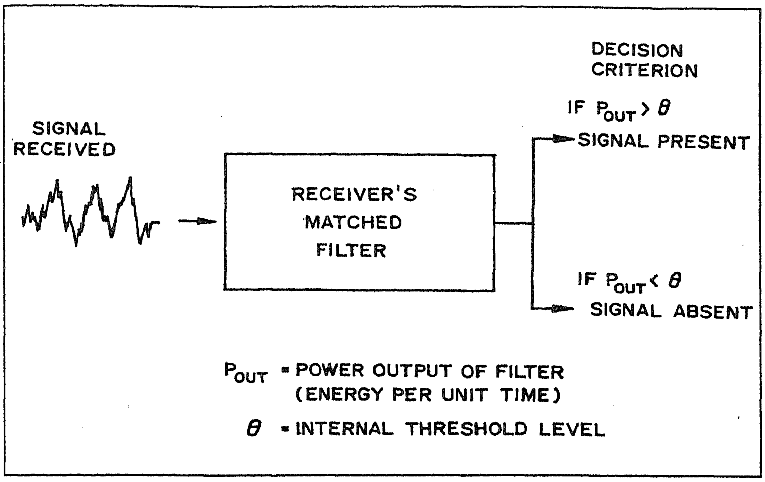


Figure 10 Decision criteria of a matched filter detector. The received signals pass through the receiver's matched filter. Whenever the power output of the filter exceeds a certain threshold, the receiver decides that the sender's signal is present. As long as the power output remains below threshold, the receiver concludes that no signal has been sent. By adjusting the internal threshold, the reliability of the receiver's decision can be altered.

domain (Hz), presence and absence of harmonics, dominant frequency (Hz). Shift in the temporal pattern occurs to avoid acoustic interferences and increase the attractiveness of the call. The spectral pattern of the conspecific calls remain unaltered. The sounds or calls of frogs can be classified according to the context in which they are produced.

The broadcasted sounds closely match the receptor sensitivities of both the males and the females of the same species. The sound characteristics provide the frogs with an effective signalling device, one that is finely tuned to the sensor capabilities of the intended receiver, yet can be adjusted to transmission of messages (*Figure 10*). Two such adjustments are changed intensity and modulation. A slight elevation of call rate towards a full aggressive call, alerts the intruding male of his presence and if the intruder does not depart, the call is further adjusted (*Figure 11*). Similarly in dense choruses males shift the timing of their call to avoid overlap with their nearest neighbours and increase the call rate and the number of call notes to improve attractiveness. Call frequency is also size dependent, with larger individuals producing lower call frequency. In most instances, females are attracted to the lower frequency calls. In others males calling on the greatest number of nights have higher success rates. Thus

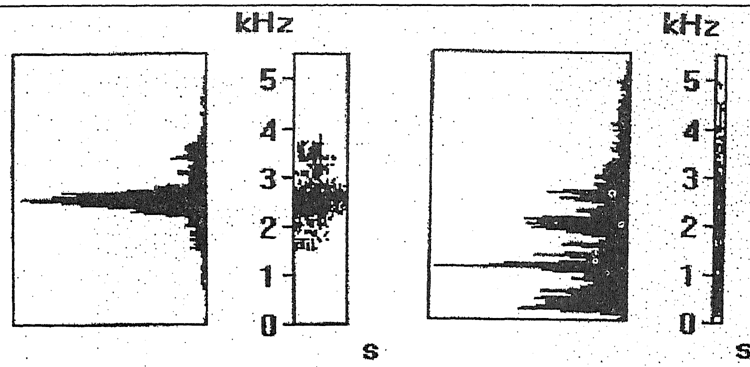


Figure 11 Male (left) and female (right) Rana erythraea [to the left]. ***Mean spectrum and sonagram of the male advertisement call of R.erythraea (left) and mean spectrum and sonagram of the female reciprocal call of R.erythraea (right)*** [*above*].

males which are larger, stronger and more durable in their nocturnal calls are favoured by the females for courtship followed by mating.

Courtship encompasses all reproductive activities prior to mating. Communication plays a key role in courtship behaviour. In the case of frogs, vocal communication plays a significant role in attracting and assessing mates. Signals help in the assessment of sex, species identity and genetic quality of the potential mate. If the potential mate passes the review, signals may increase the partner's reproductive readiness leading to the act of mating and transfer of gametes. Competition to access the female is intense, the intensity is compounded by the females selectively which compels the males to have bright colour, ornate structures and courtship behaviour to attract and retain the females. The males must therefore demonstrate their relative superiority over other males by having a stronger voice, a high quality territory or more aggressive display. Of course the qualities sought by the female frogs are not always evident to the human eye.

In 1995 it was shown that the feeble female reciprocal call given in response to the male advertisement call plays a significant role in the courtship and breeding biology of frogs.

Frog courtship is dominated by auditory signals which were thought to be the male vocalization – the advertisement call given by the males. But in 1995 it was shown that the feeble

Box 2

1. *Mating or advertisement call*: These calls attract females to the breeding sites and announce to other males that a given territory is occupied. Advertisement calls are species specific and any one species has a limited repertoire. They may also help induce psychological and physiological readiness to breed. The intensity of the call varies from species to species. The intensity of the advertisement call increases by almost 10 dB after the appearance of the female. Mating calls are emitted by male frogs.
2. *Territorial call*: These calls by male frogs may be of long or short range. They serve to demarcate and defend the territory and are emitted either sporadically or at faster call repetition rates. They may be of functional significance in the maintenance of territories or the regulation of population densities.
3. *Release call*: These are short explosive sounds repeated at irregular intervals. They often resemble an accelerated or imperfect mating call interspersed by sounds or short durations in some instances. These calls inform the partner that a frog is incapable of reproducing. They are given by unreceptive females during attempts at amplexus by a male or by males that have been mistakenly identified as females by another male.
4. *Reciprocal call*: These low pitched feeble calls are produced by females in response to the male advertisement call. Due to the feebleness of the call, it had till now escaped the attention of researchers. Recently it has been shown that only after the female produces the reciprocal call, are the final mating and egg laying activities initiated.
5. *Distress call*: These are low pitched, to shrill cries or screams. They are not associated with reproduction but are produced by either sex in response to pain or while being seized by a predator. They may be loud enough to cause a predator to release the frog.

female reciprocal call given in response to the male advertisement call plays a significant role in the courtship and breeding biology of frogs. The female reciprocal calls seem to act as a *catalyst* for the enhancement of the reproductive activity of the breeding colony. Once the female responds to the advertising males, more activity is observed in the breeding colony, involving mostly jumping around and across the responding female.

Comparative Fourier Analysis of the female reciprocal call and the conspecific male advertisement call showed that the frequency domain of the male call is almost double that of the female call and accordingly there is a shift in the dominant frequency,

Although auditory signals dominate frog courtship, the tactile and visual signals serve in the final approach and amplexus.

whereas the spectral pattern is common to both (*Figure 11*).

Although auditory signals dominate frog courtship, the tactile and visual signals serve in the final approach and amplexus. For most frogs, the tactile role in amplexus must be emphasized as it stimulates ovulation in some frogs and oviposition in all. Migration is a common feature in the life cycle of terrestrial amphibians, due to the requirement of water for their eggs and larvae. This migration occurs just before or during courtship. Pond breeding amphibians move from their terrestrial or arboreal homes to temporary or permanent ponds. Usually the males precede the females and arrive hours or days before the females. Male frog calls or chorus guide the females to the breeding areas. They partition their breeding sites, each species having a calling microhabitat. Each male defends its own territory mostly by vocalization and when that fails, they defend by head butting, wrestling or biting.

Most frogs are twilight or nocturnal creatures. They leave their hidden recesses at dusk, based on a specific brightness level. During the day they sleep directly in the sunlight above the water. Bright daylight has a striking effect on mating behaviour in species which leave their homes only at night. Bright light and water acting together release the spawns, which are then fertilized by the amplexing male.

Suggested Reading

- R Llinnas and W Precht (Ed). *Frog Neurobiology a Handbook*. Springer-Verlag. Berlin, 1976.
- M J Ryan. *The Tungara Frog. A Study in Sexual Selection and Communication*. Univ. of Chicago Press. Chicago, 1985.
- W E Duellman and L Trueb. *Biology of Amphibians*. McGraw-Hill Book Company. New York, 1986.
- B Fritz, M J Ryan, W Wilczynski, T Hetherington and W Walkowiak (Ed). *The Evolution of Amphibian Auditory System*. John Wiley and Sons.New York, 1988.
- D Roy, B Borah and A Sarma. *Analysis and significance of female reciprocal calls in frogs*. Current Science. Vol 29 (3). pp 265-270, 1995.

Address for Correspondence
Debjani Roy
Institute of Self Organising Systems and Biophysics
North Eastern Hill University
Shillong 793022, India
Fax: (091) 364 222922

A Simple Experiment to Study the Statistical Properties of a Molecular Assembly with Two or Three State Dynamics

Harjinder Singh

A dice throwing game is used to develop the key concepts in statistical thermodynamics of systems with two or three energy states. The experiment will demonstrate how, according to the second law of thermodynamics, an increasing number of molecules in an assembly produces a particular distribution of microstates. It will also lead to a statistical definition of temperature.

Harjinder Singh is with the Department of Chemistry, Panjab University. After obtaining his master's degree from IIT Kanpur he received his doctorate from Princeton University, USA. His interests include creative writing (mostly in Hindi). He was a UGC teacher fellow with the organization EKLAVYA in M.P. for a couple of years.

Consider a beaker of water – an assembly of some 10^{24} molecules. Are these molecules identical? Are they all moving in the same direction with the same speed? Are they all occupying one particular molecular energy state?

The answer to all these questions is no. Yet, at the macroscopic level, the thermodynamic properties of this system are fairly well defined. For example, it has a fairly uniform temperature. But do all the molecules have the same temperature? What does temperature mean for a single molecule?

To investigate why in spite of utter chaos at the molecular level we still end up with well defined properties for the system as a whole, we use concepts of statistics. Elementary techniques of statistical mechanics applied to systems in equilibrium are included in all undergraduate texts of physical chemistry (See Atkin's book in Suggested Reading). We discuss here a simple classroom experiment to supplement the textbook material and clarify some of the questions which frequently arise.

Box 1

What is a Micro-state?

A micro-state is a description of a system in terms of a dynamical configuration of individual molecules or particles, such that the larger macro system obeys certain constraints. For instance, a system could be described in terms of the occupations of molecular energy states of its particles, or of the momenta and position coordinates of its particles. Suppose we have a system of free particles in which each particle has a unit mass and the system has a total energy of 100 units. Then all possible states are allowed such that the sum of the squares of the momenta of individual particles taken together is 200 units. The particles may be anywhere in space. The configuration of momenta and position of the particles at any particular time is a micro-state of the system. The constraint here is the constant total energy.

The objectives of the experiment are as follows: i) to demonstrate that as the number of molecules in a system increases, a particular distribution of micro-states (see *Box 1* for definitions) becomes overwhelmingly prominent (the values of measurements taken in a laboratory will essentially correspond to this distribution); ii) to confirm that the above observation follows from the second law of thermodynamics; and iii) to obtain a statistical definition of *temperature*.

Experiment

The experiment itself is quite simple. In fact, it was inspired by an innovative middle school exercise on probability in the Hoshangabad Science Teaching Program in Madhya Pradesh (see *Bal Vaijnanik* in Suggested Reading). It is a game of throwing dice, where the value of each throw of one die represents the state occupied by an individual molecule. Let us say we have a class of students. For a large class, the students can be divided into groups to work together on the experiment. But the larger the number of experimenters in a group, the better the results. (One can even double the sampling, with each student holding one die in the left hand and another in the right hand.) Throwing the dice determines the micro-state for a collection of N

Experiment	State I	State II
A	1, 2, 3	4, 5, 6
B	1, 2, 3, 4	5, 6

Table 1 Experiments for the case where each molecule can occupy only two states; A: degenerate (equal-probability) states; B: non-degenerate states

molecules, where N is the number of members in a team. For each experiment, we can define two possible states for each molecule. For example, if the value on the die is 1, 2, or 3, we can say the molecule is in State I, and if the value is 4, 5, or 6 it is in State II. The states can be defined differently for each experiment (see *Table 1*). You can easily go further and design additional experiments with three states, which however we will not discuss here.

Each student throws a die 100 times and records the reading of the die in a table (see *Table 2*). The table shows the possibilities for the case where only two energy states are accessible to an individual molecule. Note that a die is not a molecule, but its

Table 2 A typical set of readings for the two different experiments A and B.

Move number	Readings on die	Molecular State	
		Experiment A	Experiment B
1	3	I	I
2	2	I	I
3	6	II	II
4	4	II	I
5	5	II	II
6	1	I	I
.	.	.	.
.	.	.	.
100	5	II	II

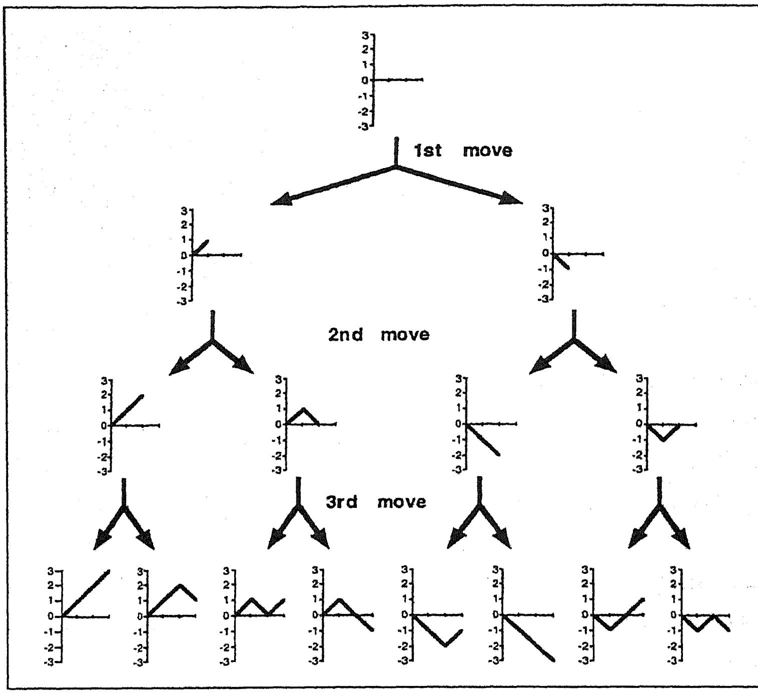


Figure 1 Possible routes of evolution of the micro-chart.

reading gives us information on the energy state occupied by the molecule. Thus, if we have just one group of ten students, we will get ten tables like the one shown representing 100 configurations (states) of ten molecules.

Although a number of experiments can be devised, *Table 1* shows only two examples. Experiments A and B represent degenerate and non-degenerate energy states respectively. In the degenerate case, both states are of equal energy and have equal probability of occupation. In the non-degenerate case, the state with a higher probability of occurrence (*which one is it?*) is the ground state, and the one with a lower probability is the excited state.

Investigating the Micro-charts: How Different are the Molecules?

For experiments A and B, each student then draws a micro-chart, as explained below and shown in *Figure 1*. We call these

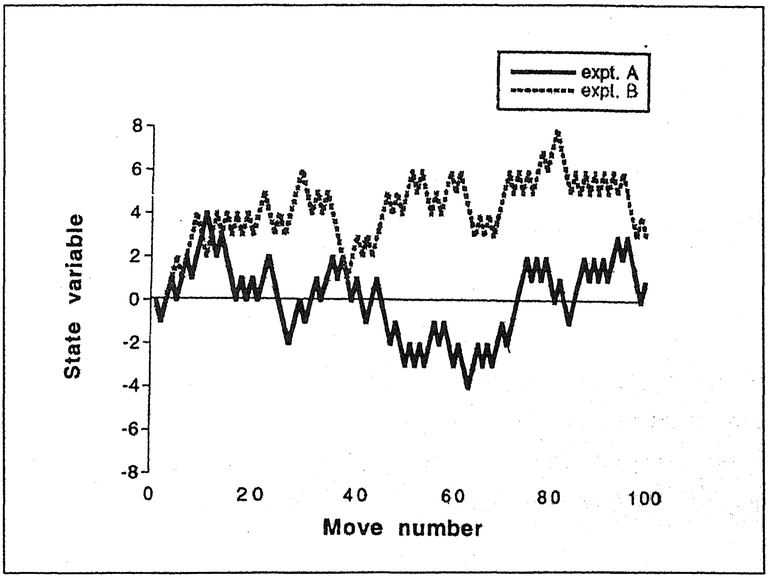


Figure 2 Typical micro-charts for the experiments A and B.

micro-charts because these depict the states occupied by individual molecules as a function of time. Starting from the origin, occupation of State I in a particular move adds +1 and that of State II adds –1 to the magnitude of a random variable plotted on the y-axis. The serial number of the move is shown on the x-axis. A typical micro-chart at the end of the experiment (100 moves) is shown in *Figure 2*. Each member of the group would find his/her micro-chart to be unique and to show no similarity with the others after the first couple of moves. This is merely a reflection of the complete randomness in the microscopic dynamics.

The Macro-chart: Towards Chemically Relevant Questions

For experiments A and B, all the experimenters in a group now work together to record a frequency table. The frequency of occurrence of a given number of molecules in State I [$n_I \in (0, 1, ..., N)$, where N is the total number of molecules] is then plotted as a bar diagram (*Figure 3*). After 25, 50 and 100 moves, the following calculations are done:

(i) the probability p_i, ($i = n_{\mathrm{I}}$) for a particular value of n_{I} is calculated as:

$$p_i = \frac{f_i \ (frequency)}{N_m \ (\#\ of\ moves)}$$

(ii) the probability p (I), that a molecule occupies State I, as:

$$p(\mathrm{I}) = \frac{1}{N} \sum_i i \cdot p_i$$

(iii) A smooth plot is drawn to clearly indicate the envelope. The peak position, n_{I}^0, peak height, full width at half maximum (fwhm) and standard deviation (Δ) are found. The ratio peak height/ fwhm are related quantities. Δ is calculated as

$$\Delta = \frac{1}{N_m} \sqrt{\sum_i (i - i_o)^2 \cdot f_i}$$

The target questions are:

(a) What is the variation in the quantities calculated above as the number of moves increases from 25 to 50? Why?
(b) Why does the peak position change from experiment A to experiment B?
(c) What is the shape of the envelope after a large number of moves? Why is it like that?

The question (a) mostly addresses aspects of statistics. For instance, the standard deviation calculated for a typical set of data as shown in the macro-chart (*Figure 3*) are shown below (*Table 3*).

Table 3 Standard deviations from the macro-chart (i_0 is assumed to be equal to the mean)

Number of moves (N_m)	Mean	Δ
25	4.60	0.3622
50	4.64	0.2593
100	4.89	0.1726

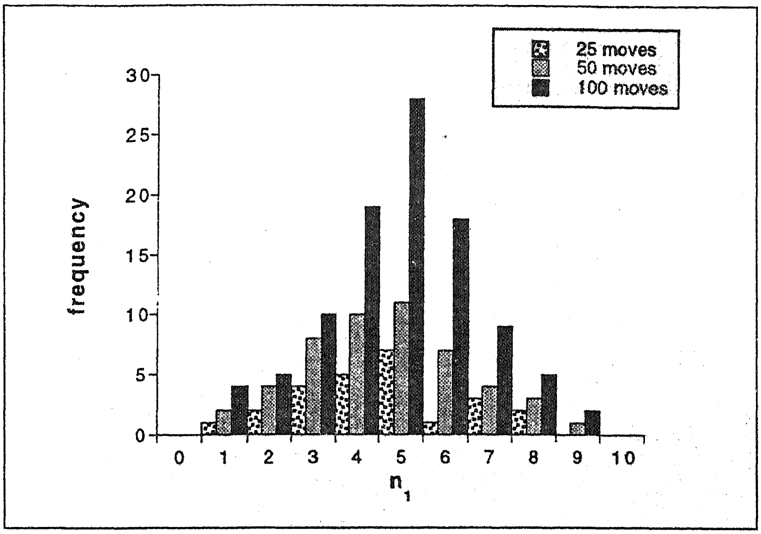


Figure 3 A typical macro-chart for experiment A.

Thus we see both visually as well as in terms of statistical calculations how the system of 10 independent molecules tends towards a most probable distribution after a large number of moves (or, in other words, a particular distribution is most likely to occur in any particular observation on a large number of molecules).

The sharpening of peak position for a particular distribution is expressed in classical thermodynamics in terms of a maximal value for the entropy, $S = k_b \cdot \ln W$, where W is the thermodynamic probability or the number of possible microstates for a particular distribution and k_b is the Boltzmann constant.

It is meaningful to calculate here the values of entropy corresponding to a couple of different distributions, including the most probable one (see *Box 2*).

Question (b) on the other hand poses an interesting learning exercise. If entropy were the only deciding factor, the peak position in the macro-charts obtained for both the experiments A and B would be the same. Is the second law of thermodynamics violated here (see *Box 3*)?

Box 2

Micro-states and Entropy

For a two state problem, with say, 10 molecules, the following table shows the values of entropy function corresponding to some of the distributions:

Number of molecules		W	S
State I	State II		cals/degree/mole
3	7	10!/(3!7!) = 120	9.513
4	6	10!/(4!6!) = 210	10.625
5	5	10!/(5!5!) = 252	10.988
6	4	10!/(4!6!) = 210	10.625

It is thus apparent that the distribution (5, 5) corresponds to the maximum value of entropy. The difference in entropy for two different distributions does not appear to be large here, but as we increase the number of molecules (to say 1000), the difference is going to increase and the value will peak sharply for the distribution with an equal number of molecules in each state (this is shown rigorously by differentiating ln W and finding the values of N_1 and N_2 corresponding to the maximum). (See McQuarrie's book in Suggested Reading)

For a three state problem the thermodynamic probability is calculated using the trinomial distribution, $W = N!/(N_1!\, N_2!\, N_3!)$.

For the non-degenerate case, since the occupation probabilities for the two different states are known (either a priori or from the frequency probabilities obtained from the macro-chart), we can now evaluate the *statistical* temperature, as

$$T^{-1} = \left(\frac{\Delta E}{R}\right) . \ln \frac{p\,(\mathrm{I})}{p\,(\mathrm{II})} .$$

This follows from the Boltzmann distribution (see*Box 4*).

Here we will talk of temperature in units of $\Delta E/R$, since we are dealing with an arbitrary separation of energy states.

Box 3

The Second Law of Thermodynamics and the Most Probable Distribution

The second law of thermodynamics states that the entropy of an isolated system tends to increase in a natural process. For a system in equilibrium, the entropy remains constant at its maximal value. For a system with a fixed composition and with only one set of degenerate energy states, interaction with the surroundings is impossible. There is no mechanism for the transfer of mass or energy; hence the system is effectively isolated. However, if there are non-degenerate states, there can be an exchange of energy with the surroundings and thus the system is not isolated any more. In such a case, the equilibrium molecular distribution depends on the collective property of thermal equilibrium. The Helmholtz free energy, $A = U-TS$ (or with appropriate conditions, the Gibbs' free energy, $G= H-TS$) determines the most probable distribution; here U is the internal energy of the system, H is the enthalpy, and T is the temperature parameter (quantifying the property of thermal equilibrium, as stated in the zeroth law of thermodynamics).

In the case of experiment B, the most probable distribution corresponds to a minimum value of A (or G). This is only valid if the macroscopic system at thermal equilibrium is constrained to have a constant volume (or pressure).

Question (c) aims at introducing the Gaussian distribution. At this juncture, another exercise may be introduced. The students may be asked to draw bar diagrams for distributions showing the number of particles in one of the two a priori equiprobable states for 2, 3, 4, 5, . . . particles. It can be easily seen that these distributions are the coefficients in a binomial expansion (the W's discussed above in *Box 2*). In the limit of large number of particles, the binomial distribution becomes a Gaussian (see McQuarrie's book in Suggested Reading).

Further Experiments

Further experiments can be designed as variations of experiments A and B. For example, a three state problem can be designed with either degenerate (e.g., I: 1, 2; II: 3, 4; III: 5, 6) or non-degenerate (e.g., I: 1, 2, 3; II: 4, 5; III: 6) states. Micro-charts for

Box 4

The Boltzmann Distribution

Perhaps the most fundamental concept in equilibrium thermodynamics, the Boltzmann distribution provides a statistical definition of temperature.

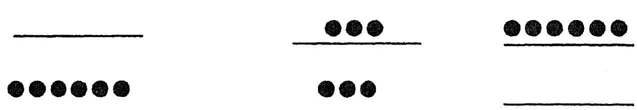
If the probability of occupation of two states with different energies are $p(\mathrm{I})$ and $p(\mathrm{II})$, then

$$\frac{\Delta E}{RT} = \ln \frac{p(\mathrm{I})}{p(\mathrm{II})}$$

where ΔE is the energy difference between the two states (*Figure 4*). The temperature, T, provides a measure of the thermal energy, $k_B T$ available to each molecule (or RT to each mole).

This holds strictly for equilibrium. In a given situation, the prevalence of equilibrium in one degree of freedom, say, translational (the velocity distribution), does not necessarily imply equilibrium in another mode. There can be different values of T for the vibrational and rotational states, each being in a local thermodynamic equilibrium.

Using the Boltzmann distribution in a nonequilibrium situation can give rise to anomalies, as for instance, in the case of population inversion (e.g., for lasers). Given below is an exercise. Find the temperature corresponding to the following distributions:



the individual molecules can be drawn with the values –1, 0 and 1 assigned to a random variable corresponding to the occupation of states I, II and III respectively. This provides qualitative information similar to that of the micro-charts from experiments A and B, on randomness at the micro-level.

One could improvise on the macro-charts and draw a 3-D plot. Obviously, this is sensible (timewise) only if a computer is accessible. Otherwise, the probability of being in a particular state is calculated directly from the bar diagram after 25, 50 and

100 moves and the results are compared with the theoretical probability. Finally, the quantity $\Delta E/RT$ is calculated for the different energy states (I and II, I and III) using the Boltzmann distribution.

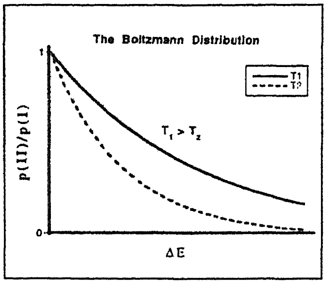


Figure 4 The Boltzmann distribution of populations in two energy states.

Conclusion

The purpose of designing this experiment was to demonstrate how the macroscopic properties of a system are seen as statistical outputs of the collective behaviour of random microscopic structures. We have avoided the use of a computer, though its use to simulate much larger numbers of dice throws would enhance the pedagogical value of the experiment and hence, when accessible, it is highly recommended (see Loetz article in Suggested Reading). In the HSTP exercise, the macro-charts are prepared teamwise as well as for the whole class. The difference in going from a team of a few members to an entire class is quite remarkable.

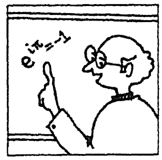
Further exercises involving the distribution function can be designed. A number of mathematical concepts, e.g., the algebra of the random walk problem as well as further questioning in the physics of the problem, e.g., the concept of phase space and the ensemble of systems with identical macroscopic properties and different microscopic descriptions, can be developed. The subject of fluctuations and nonequilibrium populations can also be invoked in an elementary manner while discussing the stochastic origins of the equilibrium distribution.

Suggested Reading

- **D McQuarrie. *Statistical Mechanics*. Harper Collins. NY, 1983.**
- **F Reif. *Fundamentals of Statistical and Thermal Physics*. McGraw Hill. NY, 1988.**
- **Bal Vaijnanik. Science text for class VIII. Hoshangabad Science Teaching Program. Madhya Pradesh Pathya Pustak Nigam, 1989.**
- **P W Atkins. *Physical Chemistry*. 5th ed., Oxford, 1994.**
- **A Loetz. *J. Chem. Ed.* Vol 72. p 128, 1995.**

Address for Correspondence
Harjinder Singh
Department of Chemistry
Panjab University
Chandigarh 160014, India

Classroom



In this section of* Resonance, *we invite readers to pose questions likely to be raised in a classroom situation. We may suggest strategies for dealing with them, or invite responses, or both. "Classroom" is equally a forum for raising broader issues and sharing personal experiences and viewpoints on matters related to teaching and learning science.

Shailesh A Shirali, Rishi Valley School, A.P.

An Elementary Problem That Interested Ramanujan!

Ramanujan used to contribute problems/ solutions to the Elementary Problems and Solutions section of the Journal of the Indian Mathematical Society. The following problem appeared in JIMS (Q 289, III 90) and Ramanujan contributed a solution JIMS(IV, 226): to determine the value of the "infinite continued square root"

$$\sqrt{1+2\sqrt{1+3\sqrt{1+4\sqrt{1+5\sqrt{1+\cdots}}}}}.$$

We proceed as follows.[1] For $x \geq 0$, define $f(x)$ by

$$f(x) = \sqrt{1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{1+\cdots}}}}.$$

To start with, suppose that $x \geq 1$. Observe that

$$f(x) = \sqrt{1+xf(x+1)}$$
$$\leq \sqrt{(1+x)f(x+1)} \leq \cdots \leq \prod_{i=1}^{\infty}(x+i)^{1/2^i}.$$

[1] The solution given here is different from that given by Srinivasa Ramanujan and follows that given by D J Newman for the same problem, in his book *A Problem Seminar*.

Using the inequalities $x + i \le 2xi$ for $x, i \ge 1$ and $\prod i^{1/2^i} < \prod 2^{(i-1)/2^i} = 2$, we obtain

$$\prod_{i=1}^{\infty}(x+i)^{1/2^i} \le$$

$$\prod_{i=1}^{\infty}(2xi)^{1/2^i} = 2x \cdot \left(\prod_{i=1}^{\infty} i^{1/2^i}\right) < 2x \cdot 2 = 4x\,.$$

Thus $f(x) < 4x$ for $x \ge 1$. If $0 < x < 1$ then $x+1 > 1$, leading to $f(x+1) < 4(x+1)$, and

$$f(x) < \sqrt{1 + x(4x+4)} = 1 + 2x < 1 + 4x\,.$$

It follows that $f(x) \le 1 + 4x$ for all $x \ge 0$.

The iterative dance now commences. Suppose that for some $a > 0$ we know that $f(x) \le 1 + ax$ for all $x \ge 0$. Then $f(x+1) \le 1 + a + ax$, and so

$$\begin{aligned} f(x) &\le \sqrt{1 + x(1 + a + ax)} \\ &= \sqrt{1 + (a+1)x + ax^2} \\ &\le 1 + \frac{a+1}{2}x, \end{aligned}$$

since $a \le \left(\frac{a+1}{2}\right)^2$.

Thus $f(x) \le 1 + ax$ implies $f(x) \le 1 + \frac{a+1}{2}x$. Repeating this step, and noting that for any $a > 0$ the sequence a, $\frac{a+1}{2}$, $\frac{a+3}{4}$, $\frac{a+7}{8}, \ldots$ converges to 1, we conclude that $f(x) \le 1 + x$ for all $x \ge 0$.

Now for the lower bound. Clearly $f(x+1) \ge f(x)$ for all $x \ge 0$, so it follows that $f(x) \ge \sqrt{1 + xf(x)}$. The

inequality is easily solved to give $f(x) \geq (x + \sqrt{x^2+4})/2$. We therefore have,

$$f(x) \geq \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}} \geq 1 + \frac{x}{2} \text{ for all } x \geq 0\,.$$

The anchor has been secured and the iteration can begin. Assume that for some $a > 0$, we have $f(x) \geq 1 + ax$ for all $x \geq 0$. We now have,

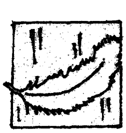
$$\begin{aligned} f(x) &= \sqrt{1 + xf(x+1)} \\ &\geq \sqrt{1 + x(1 + a + ax)} = \sqrt{1 + (1+a)x + ax^2} \\ &\geq 1 + \sqrt{a}x, \text{ since } 1 + a \geq 2\sqrt{a}\,. \end{aligned}$$

Thus $f(x) \geq 1 + ax$ implies $f(x) \geq 1 + \sqrt{a}x$. Repeating this step, and noting that for any $a > 0$, the sequence $a,\ a^{1/2},\ a^{1/4},\ a^{1/8}, \ldots$ converges to 1, we conclude that $f(x) \geq 1 + x$ for all $x \geq 0$.

The sandwich has closed with a pleasing finality about it, and we arrive at the beautiful result: $f(x) = 1 + x$. In particular, the answer to the original problem is $f(2) = 3$.

Corollary: There is precisely one function $f : [0,\ \infty) \to [0,\ \infty)$ which increases monotonically with x and for which $f(x) = \sqrt{1 + xf(x+1)}$ for all $x \geq 0$; namely, the function $f(x) = 1 + x$.

Double Your Money



Tarski belonged to a group of Polish mathematicians who frequented the celebrated Scottish Cafe in Lvov. Another member was Stefan Banach. All sorts of curious ideas came out of bull sessions in the Scottish Cafe. Among them is a theorem so ridiculous that it is almost unbelievable, known as the Banach-Tanski Paradox. It dates from 1924, and states that it is possible to dissect a solid sphere into six pieces, which can be reassembled, by rigid motions, to form two solid spheres each the same size as the original.

But what about the volume? It doubles. Surely that's impossible? The trick is that the pieces are so complicated that they don't have volumes. The total volume can change. Because the pieces are so complicated, with arbitrarily fine detail, you can't actually carry out this dissection on a lump of physical matter. A good job too, it would ruin the gold market.

The Problems of Mathematics, Ian Stewart, Oxford University Press, 1992, pp. 173.

Think It Over

This section of* Resonance *is meant to raise thought-provoking, interesting, or just plain brain-teasing questions every month, and discuss answers a few months later. Readers are welcome to send in suggestions for such questions, solutions to questions already posed, comments on the solutions discussed in the journal, etc. to* Resonance *Indian Academy of Sciences, Bangalore 560 080, with "Think It Over" written on the cover or card to help us sort the correspondence. Due to limitations of space, it may not be possible to use all the material received. However, the coordinators of this section (currently A Sitaram and R Nityananda) will try and select items which best illustrate various ideas and concepts, for inclusion in this section.

What's in a Name?

J Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560012, India

The article on unusual isomers of benzene featured in *Resonance*, Vol. 1, No. 2 has elicited many interesting responses from our readers. Quite a few have pointed out that isomer 2 with two triple bonds was incorrectly called dimethylbutadiene (instead of *diyne*). Some have ventured to name the isomer following the rules of the International Union of Pure and Applied Chemistry, but were not successful in the process. The correct IUPAC name is 2, 4-hexadiyne.

Professor Manfred Christl has written that he was quite pleased that the isomers made by his group were named after him in the article. But he wonders whether *Shakespeare benzenes* should be called *Johnson benzenes*. Although William C Shakespeare is the first author of the publication on these isomers, Professor Richard P Johnson is the senior author. Given two possibilities, I used the name that is likely to catch the attention of students and teachers. But there is nothing official about it! After all, *Shakespeare*

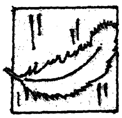
benzenes by any other name would be just as interesting, from the chemical point of view.

J Jebakumar, Lecturer at Madras Christian College, has considered the stability and possible formation of other isomers related to those discussed in the article. In particular, he has suggested that the variant cyclohex-1-ene-4-yne may be more stable than the isomers mentioned. High level quantum chemical calculations are often very useful in predicting the relative energies and electronic structures of small molecules of this type. For the interested readers, some important references to experimental and theoretical work on these isomers are given below.

Suggested Reading

- W C Shakespeare and R P Johnson. *J. Am. Chem. Soc.* Vol 112. pp 8578–8579, 1990.
- M Christl, M Braun and G Muller. *Angew. Chem. Int. Ed. (Engl.)*. Vol 31. pp 473–476, 1992.
- R Janoschek. *Angew. Chem. Int. Ed. (Engl.)*. Vol 31. pp 476–478, 1992.

Taylor Series - a Matter of Life or Death



Mathematics can even be a matter of life or death. During the Russian revolution, the mathematical physicist Igor Tamm was seized by anti-communist vigilantes at a village near Odessa where he had gone to barter for food. They suspected he was an anti-Ukrainian communist agitator and dragged him off to their leader.

Asked what he did for a living he said that he was a mathematician. The sceptical gang-leader began to finger the bullets and grenades slung around his neck. "All right", he said, "calculate the error when the Taylor series approximation of a function is truncated after n terms. Do this and you will go free; fail and you will be shot". Tamm slowly calculated the answer in the dust with his quivering finger. When he had finished the bandit cast his eye over the answer and waved him on his way.

Tamm won the 1958 Nobel prize for Physics but he never did discover the identity of the unusual bandit leader. But he found a sure way to concentrate his students' minds on the practical importance of Mathematics.

From The Observer (UK)

The Nobel Prize In Physics – 1996

R Srinivasan

The Nobel prize in Physics for 1996 has been awarded to Lee, Osheroff, and Richardson, working in Cornell, for the discovery of superfluid phases in ^{3}He, an isotope of helium, in 1972. This discovery confirmed earlier theoretical speculation for the existence of the superfluid state in this isotope of helium and also brought to light many new phenomena.

Superfluidity is the phenomenon in which, under certain conditions, a liquid can flow through narrow channels without viscous resistance. This phenomenon was first discovered in liquid ^{4}He, when it was cooled below 2.17 K, by Keesom in 1927. Kapitza was awarded the Nobel prize for his work on superfluidity in liquid ^{4}He. One can understand the behaviour of the superfluid state by assuming that the liquid is made of two components, the normal component with a density ρ_n, which has normal viscosity, and a superfluid component with a density ρ_s which has no viscosity. The total density of the liquid is the sum of the densities of the normal and superfluid components with the ratio ρ_s/ρ, where ρ is the total density, increasing from zero at 2.17 K to unity as the temperature approaches absolute zero. ^{4}He nucleus has two protons and two neutrons and its total spin is zero. A nucleus with zero or integral spin obeys Bose-Einstein (BE) statistics in which there is no restriction on the number of particles occupying a given quantum state. One of the consequences of this condition is that, below a temperature dependent on the number density of the particles, a macroscopic fraction of the total number of particles start occupying the ground state. This is called Bose-Einstein condensation (see *Resonance* in Suggested Reading). In the condensed state the particles have zero entropy. The superfluid component in ^{4}He is identified with this condensed state.

Helium has an isotope, ^{3}He, which has two protons and one neutron. This nucleus has a total spin of 1/2. All particles having half-integral spin obey the Fermi-Dirac (FD) statistics and are referred to as Fermions. Electrons in metals are characteristic examples of such particles. In the FD statistics each quantum state can be either unoccupied or can be occupied by only one particle. One would not expect particles obeying FD statistics to show superfluidity since there can be no condensation of the particles in the ground state. However some metals show superconductivity which is the resistanceless flow of electrons. Bardeen, Cooper and Schreiffer (BCS) pointed out that if two electrons have a weak attractive interaction between them, arising out of lattice deformation in their presence, then they may pair together with a total zero linear momentum and with their spins aligned in opposite directions. Since the total spin of the bound pair is zero, the angular

momentum quantum number of the pair can only be an even number. In many metals this is zero. Such paired electrons can then condense into a state in which they will show no frictional resistance to motion.

The isotope ^{3}He has an extremely low abundance in naturally occuring helium gas and so it has to be produced artifically in a nuclear reactor. This isotope became available in sufficient quantities by 1958 and extensive experiments were done on the properties of liquid ^{3}He till about 1970. These experiments established that liquid ^{3}He behaved like a normal Fermi liquid(FL)(ie) a liquid containing interacting Fermions. The interaction between the particles arises from spin fluctuations and is attractive at large distances and strongly repulsive at short distances. It is obvious that if a BCS like pairing should occur between two ^{3}He particles, they should be at large distances. The only way this is possible is for the two particles to rotate around each other with large enough angular momentum in the attractive centripetal force between the two particles. Also because of the magnetic interaction between the nuclei, in liquid ^{3}He, there are clusters of atoms with parallel spins. Pitaevskii in 1959, and other authors in 1960, predicted that one could have a pairing in which the two nuclei will have parallel spins. The total spin of the pair becomes 1 and the pair will obey BE statistics. Liquid ^{3}He should exhibit interesting anisotropic superfluid properties at low enough temperatures. The estimates of the temperature for the onset of superfluid behaviour were always lower than the lowest temperature attainable with the existing cryogenic techniques at that time.

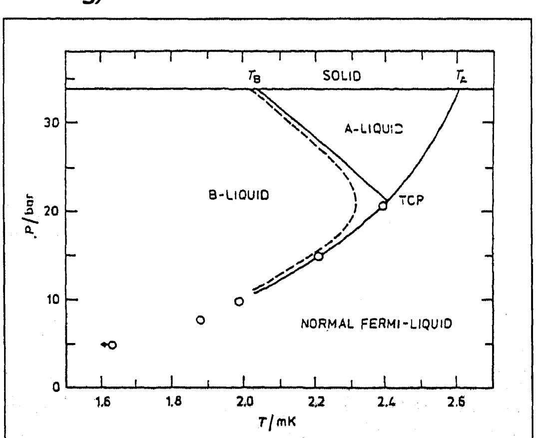
But with the development of the dilution refrigerator, Pomeranchuk cooling and adiabatic nuclear demagnetization (techniques to be described in Parts III and IV of the series article titled *The approach to absolute zero* in this journal), it became possible in 1972 to study the behaviour of ^{3}He at temperatures a few millikelvin above absolute zero.

In a Pomeranchuk cell one applies a high pressure (about 33.5 bars) on liquid ^{3}He when its temperature is a few millikelvin. When such a pressure is applied, under adiabatic conditions, on liquid ^{3}He at a temperature below 100 mK, the liquid cools with the gradual conversion of the liquid into solid. Osheroff, Richardson and Lee were looking at the pressure on the liquid as a function of time as the liquid was compressed. They observed two very small but reproducible effects at 2.6 mK and 2 mK. At the higher temperature the slope of pressure versus time graph changed abruptly, while at the lower temperature of 2 mK there was a sudden discontinuous fall in pressure after which the pressure again continued increasing. At first they thought that these changes were associated with some transitions in the solid ^{3}He phase. But NMR experiments, in which they were able to look at the NMR signal from different sections of the cell with

a technique similar to magnetic resonance imaging, clearly showed that these changes were occurring in the liquid and not in the solid. It immediately occurred to them and their co-workers that these may reflect transitions from the normal to the superfluid states expected by theorists since 1959.

Three groups, one at Cornell, the other at San Diego, USA, and the third at Helsinki, Finland, started intensive studies on various properties of liquid ^{3}He in these states. Within two years they were able to establish the phase diagram of the transition from a ^{3}He FL phase to a phase called A and then to a phase called B (see *Figure 1*) as a function of pressure, temperature and magnetic field. In the presence of a magnetic field the temperature for the A-B transition decreases. Also there is no tricritical point (TCP) at

Figure 1 The phase diagram of ^{3}He below 3mK. Dashed line shows the boundary between the A - and B- liquids in an external magnetic field of 38mT (see Lounasmaa in Suggested Reading).



which the three phases can co-exist. In the presence of a magnetic field there is a narrow region of A phase between B and FL phases. The A phase itself splits into two regions A1 and A2. When the two nuclei pair to have a spin $S = 1$, there are three possible values of M_s, namely 1,0 and −1 . There is evidence to show that phase A consists of pairs with M_s = 1 and −1 (Anderson-Morel state) and phase B contains M_s = 0 pairs also (Balian-Werthamer state).

The three main experiments to clearly demonstrate the superfluidity of A and B phases were (a) measurement of specific heat jump at the transition (b) the velocity of fourth sound below the transition and (c) direct determination of viscosity through the transition. The first measurement showed that there is a sudden jump in the specific heat at the transition FL-A, exactly what one sees in a normal to superconducting transition in a metal. When a superfluid fills a porous medium with narrow pores, the normal component is clamped because its viscosity is high. However the superfluid component can carry a pressure wave. This is called fourth sound. The velocity of fourth sound is different from the velocity of ordinary (or first) sound. There will be no fourth sound propagation in the absence of a superfluid. The density of the superfluid component is proportional to the square of the fourth sound velocity and was found to increase linearly with decreasing temperature below FL to A transition. Finally the viscosity was measured by a vibrating wire technique. The square of

the amplitude of vibration at resonance is inversely proportional to the viscosity. It was found that as the temperature passes through FL-A transition, there is a small fall in viscosity. But at the A-B transition the viscosity drops discontinuously and continues to fall as the temperature is reduced. The measured viscosity entirely arises from the normal component, the density of which decreases as one lowers the temperature below the superfluid transition.

Since the discovery of superfluidity in ^{3}He considerable work has been done over the last twenty five years, on elucidating the properties of superfluid helium, the effect of magnetic field on the phases and on the flow properties of the liquid.

One of the fall-outs of these investigations is the realization that one may have in certain metals a pairing of electrons similar to that in ^{3}He (ie) p wave pairing. Some compounds containing the heavy rare earth elements or the actinide elements, such as $CeCu_2Si_2$, UPt_3, become superconductors at very low temperatures. In these materials specific heat and magnetic susceptibility measurements indicate that the effective mass of the electrons is very high, of the order of a few hundred electron masses. These materials show unusual superconducting properties. They are called Heavy Fermion Superconductors. It is believed that the pairing of electrons in these materials is like the pairing in liquid ^{3}He .

The discovery of superfluidity in ^{3}He is a classic example of theory anticipating experiment, the development of techniques and ingenuity in experimentation going hand in hand to carry out successfully difficult experiments at very low temperatures, and the experiments in turn leading to frenetic theoretical activity which has wide repercussion in other areas of work.

Suggested Reading

- **O V Lounasmaa. Experimental Principles and Methods below 1 K. *Contemporary Physics*. Vol 15. p 353, 1974.**
- **R Nityananda. Research News. *Resonance*. Vol. 1. No. 2. p 111, 1996.**

R Srinivasan is with Raman Research Institute, C V Raman Avenue, Sadashivanagar, Bangalore 560080, India. email: rsv@rri.ernet.in

Can We Pop a Pill to Cure Obesity?

Genetic Studies of a Complex Trait

Sujatha Byravan

> When both copies of the normal *obese* gene are mutated as in the (*ob/ob*) mouse it causes profound obesity and a form of diabetes, often seen in humans.

Characteristics like obesity and many other behavioural traits are the result of the interaction of various environmental factors and lifestyle with multiple genes. This article describes recent research on the molecular biology of some forms of obesity that suggests possible ways to control the condition.

Is obesity merely the result of eating too much or is it possible that there is a genetic basis for the condition? Scientists have been asking this question for many years and the answer that emerges from recent work is: probably both. Over the years at least five independent single-gene mutations leading to the obese phenotype have been isolated. Obesity[1] is therefore a condition that involves the expression of multiple genes. In addition to being *polygenic*, it is a complex characteristic that requires the interaction of the environment with the products of these genes. Some other traits that fall into this category include intelligence, schizophrenia and congenital heart disease.

Cloning of the *ob* Gene

A *recessive* mouse mutation called *obese (ob)* was first identified in the 1950s. When both copies of the normal *obese* gene are mutated as in the (*ob*/*ob*) mouse it causes profound obesity and a form of diabetes, often seen in humans. Some scientists recently reported the cloning of the mouse *obese* gene (*ob*) and its human homologue in the journal *Nature*. This gene appears to play an important role in the maintenance of body weight and appetite. The *ob* gene codes for a 4.5 kb[2] mRNA (messenger, ribonucleic acid) capable of making a protein with 167 amino acids (See *Box 1*). Its coding sequence is *highly conserved* or, in other words, largely unchanged among various vertebrate species. When the amino acids of the Ob[3] protein in mice and humans are compared, 84% of them

[1] Obesity implies body weight in excess or equal to 20% above ideal body weight. Generally obesity is determined by comparing standard tables for height and weight or by calculation of body mass index (body weight in kg/height in meters). More precise indications can be obtained by skin-fold measurements.

[2] 1kb = 1000 nucleotide bases

[3] ob refers to the gene while Ob refers to the protein

[4] 1kD = 1000 dalton, 1 dalton = 1.66024×10^{-24} g.

are identical. Such a high degree of sequence similarity conserved across species suggests that the protein plays a very important role in the animal. Ob protein is a single subunit of 16 kD[4] and is synthesised mainly in the adipose tissue. It is now called *leptin*, a term derived from the Greek root *leptos*, meaning thin.

Following the cloning of the *obese* gene, scientists from Amgen Inc., Rockefeller University and Hoffmann-La Roche Inc. independently published some interesting results. They surmised that if a mouse is obese because of a mutant gene product, injections of the normal gene product might correct the condition. Indeed, all three groups showed that when the product of the normal *obese* gene was injected into a mutant mouse it induced weight loss.

Results of the Leptin Experiments

Daily injections of leptin in obese mice had two effects: reduced appetite and increased energy use. These changes led to a reduction in food intake and the percentage of body fat, along with an increase in metabolic rate, body temperature and activity levels. Surprisingly, injections of both the mouse and human leptin (10 μg/g body weight/per day) reduced body weight in obese mice compared to normal animals given a saline injection. When mice that were not obese were given leptin they too lost weight and maintained their new weight as long as they continued to receive the injections.

The results suggest that the Ob protein or leptin plays a role, maybe even a pivotal one, in the regulation of body weight and the deposition of fat. The protein may signal the brain so that if its levels are high, less food is consumed and more calories burnt, while when its levels are low and as appetite increases more food is consumed. Injections of leptin may therefore trick the body into thinking that it is satiated, thereby reducing appetite and burning extra calories (*Figure 1*).

Other Genes Regulating Obesity

For a long time it was suspected that a region of the brain called the hypothalamus may be important in regulating body weight since lesions in this area led to obesity in rats. In

Box 1

The Transfer of Genetic Information

The sequence of nucleotides in the DNA determines the sequence of ribonucleotides in the RNA. The messenger RNA (mRNA) is that class of RNA molecules which is used to synthesise proteins. The sequence of ribonucleotides is read as *triplet codons* where each codon specifies an amino acid in the protein. The synthesis of RNA molecules from the DNA is called *transcription* and the synthesis of proteins from the mRNA is referred to as *translation*. This transfer of genetic information from DNA → RNA → Protein is often called the *central dogma*.

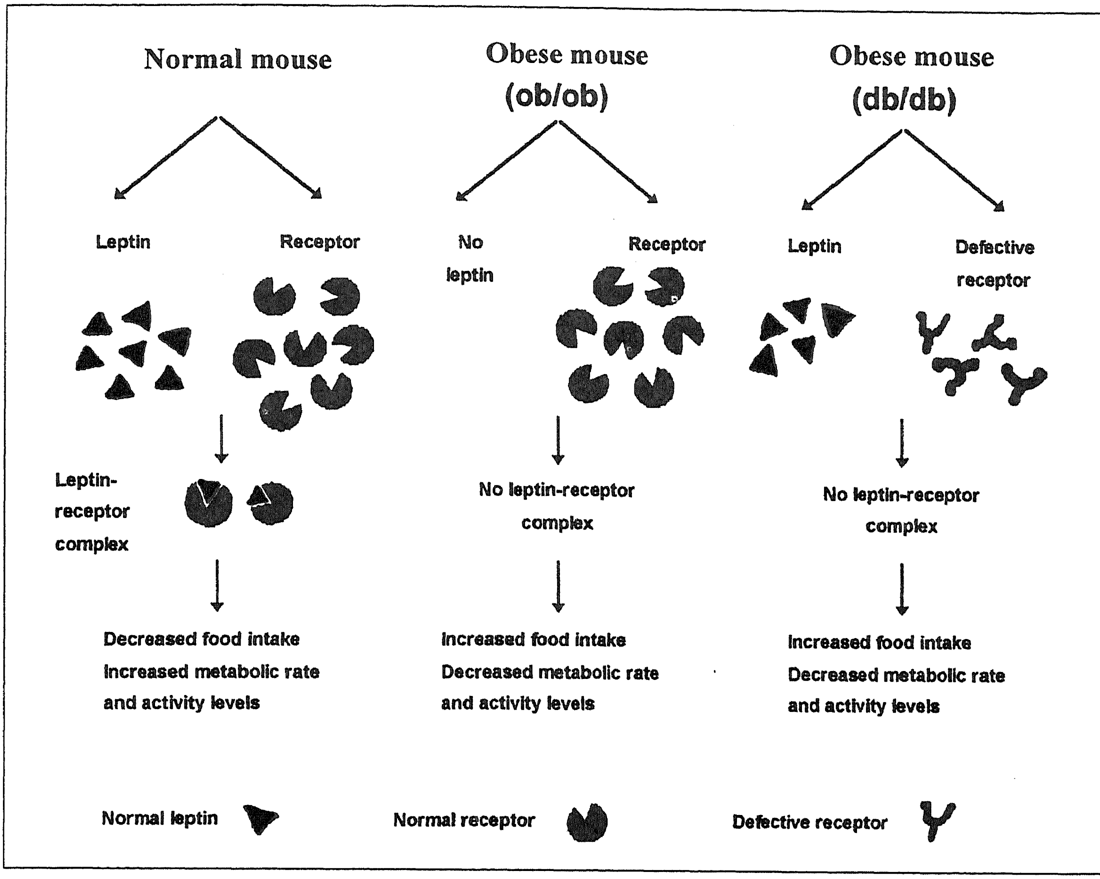


Figure 1 The leptin-receptor complex is formed in the normal mouse but not in the obese mouse with ob/ob or db/db mutations.

December 1995, a research team reported that they had found yet another mutation which resulted in obesity in spite of the presence of normal leptin. They identified the defect in this case to be in the receptor (Ob–R) that is expressed in the hypothalamus and binds to leptin. The gene for the receptor is designated as the db gene. The interaction between leptin and its receptor is essential in the signalling system that involves the hypothalamus and normally results in satiation. When this db gene is mutated, as is the case in the *db/db* mouse, a defective receptor is made and since leptin is unable to bind to it no satiation signal is sent. This consequently leads to obesity (*Figure 1*).

Another team is working on a strain of mice that grows obese when the diet contains excessive fat. This is called *diet-induced-obesity*. Yet another strain of mice stays trim in its youth but puts on weight in adulthood. This is labelled *maturity-onset-obesity*. Since these forms of obesity are similar to those in

humans, their genetic basis is being studied with interest.

The Social Context of Obesity Research

Severe weight gain is a health hazard but we see that weight loss has become a major preoccupation even with the healthy. Since many cultures place increasing emphasis on being thin weight loss products are bound to become very popular. It now appears that some scientists are about to join this profit making diet bandwagon. Amgen Inc., a biotechnology firm in California has paid $20 million to the Rockefeller University for a license to develop products using the *ob* gene. Returning to our original question, does this mean that we are closer to popping a pill to cure obesity? Not really. We know very little about the various genes that regulate body weight, their interactions, the role of the environment, and the possible side effects of popping a pill to cure the condition, even if that were possible. In fact according to Bruce Spiegelman who studies obesity at the Dana-Farber Cancer Institute in Boston, "Nobody has found or developed an organic compound that can mimic insulin, and we have had the insulin receptor in hand for years." Thus, while this finding is interesting for its own sake, at this stage it is still doubtful whether it has direct clinical applications. Perhaps in the long run those with complicated health-threatening obesity could avail of a simple remedy. Unfortunately, what appears more likely in the short-term is the commercial exploitation and consequent potential misuse of this information through the collaboration of two already well established giants: the weight loss and the biotechnology industries.

Suggested Reading

- Y Zhang *et al. Nature*. 372:425–431, 1994.
- MA Pelleymounter *et al. Science*. 269:540–543, 1995.
- JL Halaas *et al. Science*. 269:543–546, 1995.
- LA Campfield *et al. Science*. 269:546–549, 1995.
- GH Lee *et al. Nature*. 379:632–635, 1996.

Sujatha Byravan is at the Indian Academy of Sciences, Bangalore 560080, India. email: sujatha@ias.ernet.in

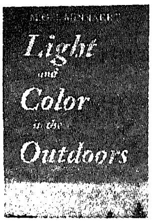
Grandpa, your blood pressure is getting high and your volume has increased a lot. Is this not a deviation from Boyle's law?

Sumanta Baruah, Class XI, Biswanath College, Assam

The Magic Of Light

A Vision of Nature in All Her Colours

G S Ranganath



Light and Color in the Outdoors
MGJ Minnaert
Springer-Verlag, 1993
pp.417, DM78

When we are face to face with the grandeur and beauty of nature we cannot remain mute spectators. Prof Marcel Gilles Jozef Minnaert was no exception. During his long and eventful career as an optical astronomer he could not help admiring the beauty of the rainbows, the glories, the coronae and so on. He studied them extensively and the book under review is a product of such a labour of love. In this book Minnaert explores the splendour and the spectacle of light in nature. The hallmark of this work is the variety and richness of natural optical phenomena that are described and discussed. More than 250 optical effects have been covered which remind us of the famous statement by Sherlock Holmes "we see but do not observe". One who reads this book will be tempted to look more closely and carefully at all that has only been casually seen earlier. In the end one is left with the impression that nature has revealed the beauty of light while manifesting herself optically. Before I

One who reads this book will be tempted to look more closely and cqrefully at all that has only been casually seen earlier. In the end one is left with the impression that nature has revealed the beauty of light while manifesting herself optically.

undertake the exercise of justifying these statements I wish to make a few remarks. The book under review is a new and revised edition of an earlier book entitled *Light and Colour in the Open Air* published in 1940. The new edition which also happens to be a new English translation was brought out on the hundredth birth anniversary of Minnaert. I will first of all highlight features common to both these editions. Later I will emphasise what is new and interesting in the latest edition.

We have all seen our shadows in the open sunshine many times. Only a few of us would have observed that the shadow of our head is very hazy and fuzzy when the sun is low in the sky. It is also not always appreciated that the shadows of leaves, butterflies and small low flying birds are all nearly circular patches. Strangely, high flying birds do not cast shadows at all. Interestingly all these and many more equally intriguing shadows are manifestations of the sun not being a point source but a circular disk of light.

We generally take for granted that images seen in reflection are faithful to the objects

Box 1

Marcel Gilles Jozef Minnaert was born on February 12, 1893 in Bruges, Belgium. At the University of Ghent he studied biology and completed his doctoral thesis at the age of 21 on the effects of light on plants . After getting his second doctorate on anomalous dispersion, this time from the University of Utrecht, he worked at the Utrecht Observatory on solar astronomy, comets and the photometry of Venus and the Orion nebula.

It is very instructive to read the following passage from the preface to the first Dutch edition of his book *Light and Colour in the Open Air.*

"..... indeed there is hardly a branch of physics that is not applicable out of doors, and often on a scale exceeding any experiments in a laboratory. Bear in mind, therefore, that everything described in this book lies within your own powers of understanding and observation. Everything is meant to be seen by you and done by you! Where the explanations offered are perhaps too concise, I suggest that you refresh your memory of fundamental physics by turning to an appropriate elementary textbook.

The importance of outdoor observations for the teaching of physics has not yet been sufficiently realized. They help us in our ever increasing efforts to adapt our education to the requirements of everyday life; they lead us naturally to ask a thousand questions, and, thanks to them, we find later on that what we learned at school is to be found again and again beyond the school walls".

being reflected. This may not always be true. For example, when we see the reflected image of a rainbow against a patch of cloud, surprisingly we see a relative shift between the images of the cloud and the rainbow. If we look at our own shadow on a bed of dried leaves, strangely, we see halos around the shadow of our own head but not around the shadow of our neighbour's head. These and other equally tantalizing optical wonders described in this book, are elementary consequences of the laws of reflection of light under different settings.

A whole chapter is dedicated to rainbows, halos and coronae. The optics of refraction and diffraction in water droplets and ice crystals leads to a plethora of effects, the most famous being the rainbow. These are seen in a variety of situations and their discussion adds considerably to our appreciation of these beautiful optical spectacles. This is a highly educative and an instructive chapter.

Quite often our visual experiences are not the same as those seen in a photographic image of the same scene. This is largely due to the fact that our eyes do not work exactly like cameras. For instance, a faint star when

Box 2

Why do we not see the setting of stars? Generally, we do not see the setting of stars unless they are very bright like Sirius. This is because, atmospheric scattering considerably diminishes the light that can otherwise reach an observer. This effect becomes all the more important when the stars are approaching the horizon. Thus, they become too faint for our eyes to see.

looked at straight disappears from our view but the same can be seen through the corner of our eyes. Once in a while we seem to see a star wander, rather erratically, in the sky. While seeing a double star through a moving binocular, we observe a relative motion between them with the fainter star lagging behind the brighter one. All these are entirely due to the way the eye functions. Justifiably, Minnaert has elaborated this important subject over three chapters. To the non-expert this is probably the most impressive part of the book.

We could go on listing the optical mysteries of the world around us. Instead, I will now dwell upon the special features of the new edition. Apart from a good collection of new and colourful photographs, almost in every section we find an additional example to illustrate the point under discussion. It is rather surprising that many of these were not included in the first edition since they refer to phenomena that were known even then. As an example let me mention the halo sometimes seen around the sun and moon. This has a radius around 27°. It was first recorded by Scheiner and appears to have been seen only a few times in history. Interestingly, a recent analysis of this halo indicated that it might be due to cubic ice crystals. Since ice generally crystallizes in hexagonal symmetry, this suggestion was accepted with reservation. But a few years back it was experimentally demonstrated that sometimes under fast quenching, water crystallizes in the cubic form. The sky probably still holds many more secrets to be unravelled.

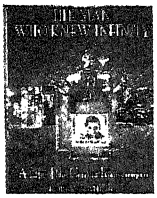
The only disappointment with both editions of this book is that once in a while it becomes a dull listing of information instead of a flow of a theme. But we should not be harsh on Minnaert who was probably burdened by the sheer weight of his observations which might have lead to some drab passages here and there. Every lover of nature, whether a student of science or a layperson, must read this book. It might be a bit expensive for an individual to possess it but is not beyond the reach of libraries.

G S Ranganath is with Raman Research Institute, Bangalore 560080, India. email: gsr@rri.ernet.in

The Man Who Knew Infinity

Untutored Genius

R Tandon



The Man Who Knew Infinity
A Life of the Genius Ramanujan
Robert Kanigel
Rupa & Co., 1992
pp.436, Rs.195

It is not very often that one sees the biography of a pure mathematician – not a mathematician like Newton who was also a physicist but a pure mathematician who worked in number theory, probably the purest realm of mathematics. Ramanujan's work may today have some applications in particle physics or in the calculation of π up to a very large number of decimal places but it would be misleading to stress this aspect of his work. Ramanujan did mathematics for its own sake, for the thrill that he got in seeing and discovering unusual relationships between various mathematical objects. The oft repeated anecdote about Ramanujan finding the number 1729 interesting because it is the smallest number that can be written as a sum of two cubes in two different ways, 10^3+9^3, 12^3+1^3 bears repetition only in that it gives a flavour of the kind of mathematics that Ramanujan was interested in – the kind of mathematics about which the sceptic will say, so what! and surely Ramanujan would have said, so what!

Ramanujan was born in a poor Tamil Brahmin family that resided in the town of Kumbakonam. He attended school there and did averagely well. While in school he came across a book entitled *A synopsis of elementary results in Pure and Applied Mathematics* by George Carr. This book is just a compendium of results on integrals, infinite series and other mathematical entities found in analysis. Yet it left a lasting impression on Rama-nujan; in fact it virtually determined his mathematical style. He would later write mathematics as a string of results without proof or with the barest outline of a proof.

After school Ramanujan was hooked on mathematics. He spent all his time with his head over a slate working on problems in number theory that interested him and neglected everything else. The result was that he could never get through another examination. An early marriage as was usual in those times led to a frantic search for a job to earn an income. He became a clerk in the Madras Port Trust with the help of some well wishers. In the meantime Ramanujan kept showing his results to various people who he thought would be interested or would help him get a job that would give him a lot of time to do mathematics. He wrote to a couple of well known British mathematicians giving a list of some of the results he had obtained. They ignored him – thought he was a crank! Finally he wrote to one of the most distinguished English mathematicians of the time – a person who had done a lot of work on

number theory – G H Hardy. Hardy arranged for Ramanujan to come to Trinity College, Cambridge where he and Ramanujan met almost daily discussing mathematics for about three years. Starting from the early life of Ramanujan, Kanigel describes all this and much more in his beautiful story on Ramanujan.

Kanigel's book is not only a biography of Ramanujan but contains a minibiography of Hardy as well. On Hardy, Kanigel is brilliant. Though basically a conservative Englishman, Hardy was a revolutionary in the world of British mathematics. (The ethos of British mathematics of the time was determined by the famous Tripos examination. It is not unusual even today to hear someone of our father's generation boasting that so and so was a Wrangler – that is had acquired a high rank in the Tripos examination. To Hardy however the Tripos was an anachronism – it had outlived its usefulness.) Kanigel writes about all this with a great deal of perception, which is especially admirable in view of the constraints he worked with. To quote Kanigel himself: "In writing the life of Ramanujan, I faced the barriers of two foreign cultures, a challenging discipline, and a distant time. As I am expert in none of these, I owe a debt of gratitude to the many persons who have helped me surmount those barriers – who have consented to interviewing, spent hours explaining recondite areas of mathematics or Indian cultural life, guided me to out of the way documents in libraries and archives, read and criticized early drafts, befriended me in England and India – and, back in Baltimore, offered a supportive hand or word of advice".

Kanigel covers in a most balanced manner two controversies that are usually associated with Ramanujan - one: whether Ramanujan drew upon divine or some sort of mystical inspiration to come up with his ideas or whether as was maintained by Hardy he was just a hard working, exceptionally original and creative person like other great mathematicians. The second controversy concerns the medical cause of Ramanujan's death.

So, what did I get from Robert Kanigel's book on Ramanujan? I think it reinforced my belief in the universality of the language of science. Here were two totally dissimilar people both culturally and temperamentally – from two totally different backgrounds. Each knew very little about aspects of the other's personal life, and when Hardy learnt later of Ramanujan's personal problems they came as a complete surprise to him. Of course, Hardy himself was too reserved to let Ramanujan get even a whiff of his own concerns. But when they met, which they did almost everyday for nearly three years, they were on exactly the same wavelength – they spoke exactly the same language – they were totally *intimate* with each other in the language of mathematics.

I would like to end by quoting from the last few pages of the book. It was 1936, Harvard University was celebrating the 300th

anniversary of its founding and as part of the celebrations it was honoring sixty eminent men of letters and science. The list included Einstein, Heisenberg, Piaget and Jung. They were also honoring Hardy. Hardy gave a series of lectures that later came out as a book entitled *Ramanujan: Twelve lectures on subjects suggested by his life and work*. Hardy said at the beginning of his lectures, and I quote from Kanigel's book:

"I have to form myself, as I have never really formed before, and to try to help you to form, some sort of reasoned estimate of the most romantic figure in the recent history of mathematics, a man whose career seems full of paradoxes and contradictions, who defies almost all canons by which we are accustomed to judge one another and about whom all of us will probably agree in one judgement only, that he was in some sense a very great mathematician".

And then Hardy began to speak about his friend Ramanujan.

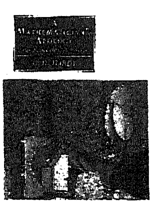
Kanigel tries to do the same. Hardy concentrated on the technical aspects of Ramanujan's work. Kanigel in his own meticulously researched style tells us more about Ramanujan the man.

R Tandon is with the Department of Mathematics, University of Hyderabad, Hyderabad 500007, India.

A Mathematician's Apology

Mathematics and Creativity

R Vittal Rao



A Mathematician's Apology
With a Foreword by C P Snow
G H Hardy
Cambridge University Press, 1993
pp153, Rs. 175

Every creative person – whether a painter, a poet or a scientist – reflects, at some time or the other in his life, on whether his work is worth doing and why he does it. It is these very questions that Hardy – one of the outstanding mathematicians of this century – takes upon himself, as a mathematician, to answer, and "put forward an apology for mathematics". The book has been reviewed, discussed and debated by several mathematicians. It is natural that the author's personal likes and dislikes play an important role in answering these questions. One must keep this in mind while reading the book. It is quite likely that a reader will get involved (as this reviewer did) in an argument and disagree with the author every now and then while reading through the book!

Another crucial aspect that a reader must keep in mind is that Hardy authored this book while he was in his sixties and felt that his creative powers were on the decline. No wonder, then, that he begins the book with the words, "It is a melancholy experience for

a professional mathematician to find himself writing about mathematics". Hardy believed in doing mathematics and not writing or talking about mathematics. C P Snow, in his foreword, aptly writes, "*A Mathematician's Apology* is, if read with the textual attention it deserves, a book of haunting sadness".

Hardy quickly tackles the question of *why he does it* by the argument that "I do what I do because it is the one and only thing that I can do at all well .... It is a tiny minority who can do anything *really* well, and the number of men who can do two things well is negligible". The sadness of his age having drained his creativity haunts him when he writes, "No mathematician should ever allow himself to forget that mathematics, more than any art or science, is a young man's game". This view has been debated for years. "I do not know an instance of a major mathematical advance initiated by a man past fifty". While one may generally agree that with the advance of age, a mathematician does lose his sharpness, his creativity, yet as another outstanding mathematician Mordell says "there are still many consolations . . . find pleasure in thinking about some of our past work . . . sometimes completely changing the exposition of classical mathematics .... (We can still be of service to younger mathematicians)".

Hardy attempts in detail to analyse the question of whether his work is worth doing. He puts forth several arguments. The harmlessness of mathematics, its *profitability* or otherwise, the permanence of mathematical achievement are some of the ideas that Hardy touches upon. To him the patterns created by a mathematician are more permanent than those of a painter or a poet since those of a mathematician are woven with *ideas*.

Hardy is very emphatic that the mathematical patterns "must be beautiful – there is no permanent place in the world for ugly mathematics" – though, it "may be very hard to *define* mathematical beauty, but this is just as true of beauty of any kind". To bring home this point, Hardy discusses some examples. He is somewhat handicapped by the fact that he must choose examples that are very simple and intelligible to a nonspecialist. He chooses to discuss the aesthetic appeal of the infinitude of primes and is thrilled by the "very high degree of *unexpectedness*, combined with *inevitability* and *economy*". Hardy is not the type who considers problems involving enumeration as *real* mathematics – "enumeration of cases, indeed, is one of the duller forms of mathematical argument". A real mathematical proof, according to Hardy, should be like a "simple and clear-cut constellation, not a scattered cluster in the Milky Way".

Hardy shows his strongest prejudices when he discusses the *usefulness* of mathematics. His views will certainly be debated for ever. Some glimpses of his views can be found in the following words: ". . . even Littlewood could not make ballistics respectable, and if he could not who can?", "The 'real'

mathematics of the real mathematicians, the mathematics of Fermat and Euler and Gauss and Abel and Riemann, is almost wholly 'useless' ". Hardy's strong dislike for *usefulness* or application of mathematics, in particular, and science in general, stems from his obsession with their possible use in war. For instance, he observes, "But science works for evil as well as for good (and particularly, of course, in time of a war); and both Gauss and lesser mathematicians may be justified in rejoicing that there is one science at any rate, and that their own, whose very remoteness from ordinary human activities should keep it gentle and clean". Should Hardy be alive today he will surely be a sad man to see the military applications of *real* mathematics of the *real* mathematicians!

Many of Hardy's views are too harsh and sweeping to be acceptable. His views on usefulness of mathematics discussed above is a good example of this. To cite just one more, he writes, "– good work is not done by *humble* men".

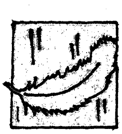
It is very hard to believe that Hardy felt, as a boy, no passion for mathematics (Section 29) and thought of mathematics more in terms of examinations and scholarships. This is in complete contradiction to the general belief that most great mathematicians are born and not made, and their interest develops and their talent shows up from a very young age.

Hardy's beautiful essay makes very interesting reading, giving glimpses of the mind of an intellectual, saddened by his loss of creativity, and proud of his being a *real* mathematician. It must be read not only by mathematicians and scientists, but also by historians, students of psychology and literature.

The book has a lengthy and very interesting foreword (which is almost a brief biography of Hardy) by C P Snow, Hardy's friend of long standing.

To sum it up, we may borrow a phrase from Hardy, and rate the book as of *Bradman Class*, and to borrow Snow's words, should be read *with the textual attention it deserves*.

R Vittal Rao is with the Department of Mathematics, Indian Institute of Science, Bangalore 560012, India.



Why does this magnificent applied science, which saves work and makes life easier, bring us so little happiness? The simple answer runs: Because we have not yet learned to make sensible use of it.

Albert Einstein

Reflections Around the Ramanujan Centenary

Atle Selberg

The Norwegian Mathematician Atle Selberg is considered to be one of the great number theorists of this century. He was awarded the Fields Medal in 1950 and is currently at the Institute for Advanced Study, Princeton, USA.

This article is based on the tape of an extemporaneous talk to a general audience at the Tata Institute of Fundamental Research, Bombay, India, given after the conclusion of the Centenary Conference held there (January 1988).[1]

Srinivasa Ramanujan's work played a very important role in my own development as a mathematician. I first saw his name in 1934, when I came across an article in the periodical of the Norwegian Mathematical Society. It was a periodical that my father subscribed to and the title of the article was, if I translate it from Norwegian, "The Indian Srinivasa Ramanujan, a remarkable mathematical genius". But I should add that, since there is no one-to-one correspondence between languages, for the word I have translated here as "remarkable", the Norwegian word carries also a connotation meaning "unusual and somewhat strange". It was written by a professor of mathematics at the University of Oslo by the name of Carl Störmer who had actually started out by being interested in Number Theory in his youth, later turned to the mathematical theory of the aurora borealis (northern lights) and made quite a name for himself in connection with that. But he retained his interest in mathematics (Pure mathematics).

The material for his article was mainly taken from the biographical articles in the *Collected Papers* of Ramanujan which had been published in 1927 by the Cambridge University Press. It gave a sketch of the history of Ramanujan's life and he quoted quite a number of his results and samples of what I thought were extremely remarkable, strange, and beautiful formulas. It was, if I remember, probably about 15 pages long and not more than 20

[1] Reproduced from *Atle Selberg, Collected Papers*. Vol I", pages 695-701, Springer Verlag, 1989.

pages. But it made a very deep and lasting impression on me and it fascinated me very much.

At that time, I was still a school boy and, for some years, I had been reading mathematics on my own in a rather unsystematic and haphazard way, in my father's mathematical library which was quite large for a private individual, I may say.

I may make a digression on that, because it shows how chance plays a role in one's life - a very great role often. I had started reading mathematics when I was about thirteen or so. I had accidentally opened a book and come across Leibnitz's series for $\pi/4$: 1-1/3+1/5-1/7 and so forth involving reciprocals of the odd integers with alternating signs. Till then, school mathematics had always bored me but this seemed such a very strange and beautiful relationship that I determined I would read that book in order to find out how this formula came about.

It is a wonder that I did not stop because the book actually started with a long chapter on the concept of the real number and with the Dedekind cut which was not the most inspiring beginning.

At any rate, when I saw the article about Ramanujan, I had already made up my mind to go into mathematics but I had no more than vague notions about what kind of mathematics. I think I was at that time mostly thinking of going into the general theory of analytic functions; more specifically, something like the Nevanlinna theory which was then very much in the foreground and which actually my oldest brother had started doing. He was then a research fellow at the university. Another brother who was already a student of mathematics for a few years was grappling with the Theory of Numbers. He had also read this article on Ramanujan, and so he borrowed from the university library Ramanujan's *Collected Papers* and later he brought the book home during a vacation.

So I got a chance to browse through it for several weeks. It

seemed quite like a revelation - a completely new world to me, quite different from any mathematics book I had ever seen - with much more appeal to the imagination, I must say. And frankly, it still seems very exciting to me and also retains that air of mystery which I felt at the time. It was really what gave the impetus which started my own mathematical work. I began on my own, experimenting with what is often referred to as q-series and identities and playing around with them.

...Ramanujan's Collected Papers... It seemed quite like a revelation - a completely new world to me, quite different from any mathematics book I had ever seen - with much more appeal to the imagination...

In the summer of 1935, as I finished the gymnasium, I wrote a manuscript - I will give the title in English, *On some arithmetical identities*. German was my best foreign language at that time and I gave this manuscript to Professor Störmer in Oslo when I started my university studies at the beginning of the fall term in 1935. He sent it to Professor G N Watson in Birmingham, England, in order that it could be refereed. He had some connections with Watson before. And Watson, after having kept it for quite some time (much too long I thought at that time but now that I myself have been a referee, I have somewhat more understanding for Watson's delay!), finally sent it back with the recommendation that it be published, which it duly was in 1936.

Watson also sent me a lot of his reprints, and particularly, reprints dealing with things from Ramanujan's *Notebooks* and *Nachlass* or posthumous papers. Among these were, for instance, reprints of Watson's papers on the mock-theta functions of orders 3 and 5 where he proves the various identities or relations that Ramanujan had stated for these functions and also showed that they had the required very precise asymptotic behaviour when one approaches the roots of unity on the unit circle. And this motivated my second paper which dealt with mock-theta functions of order 7 and established their asymptotic behaviour.

In the meantime, I had acquired my own copy of Ramanujan's *Collected Papers* which was a present from my father that I carry with me now. I also thought of studying some of the other papers - first of all, the joint paper with Hardy on the partition function

If one looks at mathematics as a body of knowledge, I think it definitely can be characterized as a science, but if one looks at the way in which it grows and accumulates, the actual doing of mathematics seems much more to be an art.

expressing the number of ways in which a number can be written as a sum of integers. The paper is entitled *Asymptotic formulae in combinatory analysis*. The other papers involved modular forms and their coefficients. But this was my first contact with the theory of modular forms and automorphic forms in general, as well as with discrete or -as we in general called it at that time - discontinuous groups and this is a theme that ever since has remained one of the main interests, perhaps the main interest in my mathematical work.

Before I go on to Ramanujan's mathematics, let me say something about mathematics or mathematicians in general, for the benefit of the non-mathematicians. Let me interject that I feel truly sorry for them. I think they are missing what I think is the most exciting and rewarding intellectual activity.

Mathematics is often compared to the arts, particularly to music, and it is true that in mathematics as in music talent may flower at an astonishingly early age, though it has to be admitted, earlier in music than in mathematics. In mathematics, aesthetic considerations, beauty, simplicity and elegance are very important as well as truth. If one looks at mathematics as a body of knowledge, I think it definitely can be characterized as a science, but if one looks at the way in which it grows and accumulates, the actual doing of mathematics seems much more to be an art. Mathematics is concerned exclusively with objects and structures which are creations of the mind, although they may be suggested by or patterned on things that are found in the so-called real world. And since it deals only with creations of the mind, it is cumulative in a way that the other sciences are not.

Mathematics does not shed the old substances as new is added, in the way the natural sciences, for instance, will do. The work of Euclid, Apollonius or Archimedes, to mention some Greek mathematicians from antiquity, is as valid today as when it was done more than two millennia ago. But while the content or substance remains, the form in which it is presented is ever

changing. What we may refer to as the landscape of mathematics may change profoundly from one generation to another, and even during shorter time spans, fundamental changes may occur.

If I consider my own time as an active mathematician, the change between how I thought in the thirties and now, of course, is tremendous. The papers that appear today could not have been understood by anyone - or atleast most of them – at that time; they often deal with concepts that did not exist then. Mathematics grows in many ways by diversification and complexification or specialization, as one subject may branch out in many directions and form separate specialities. On the other hand, we also have convergence and synthesis, simplification and even unification, as different fields of mathematics that seem distant and without any connections, may develop several bridges and eventually become closely interconnected. While Indian mathematics seems to have been built on the concept of the number from quite an early stage, Greek mathematics was built on the concept of point, line and plane and the relations between these; only after the revival of mathematics in Italy, was Western mathematics primarily based on the concept of number and remained so for several hundred years. Today, we have a mathematics which is primarily concerned with structure and relationships between structures rather than just relationships between numbers. The first beginnings of this we find around 1800 and the first breakthrough in this direction was, of course, the introduction of the abstract group concept. Now it is almost all-pervasive in the field of mathematics.

Mathematical talent is also something that occurs in many varieties. Some mathematicians are theory-builders, some are problem-solvers and some originate problems - I will not say that they create problems! They originate them. Or, they may produce the first isolated examples of new mathematical objects or relationships which later give rise to comprehensive theories. None of these different abilities or talents should be ranked

Mathematical talent is also something that occurs in many varieties. Some mathematicians are theory-builders, some are problem-solvers and some originate problems ... None of these different abilities or talents should be ranked higher than others. In the long run, they are all needed for the continued prospering of mathematics.

Ramanujan's particular talent will seem to be primarily of an algebraic and combinatorial nature... He had, on his own, acquired an extraordinary skill of manipulation of algorithms, series, continued fractions and so forth, which certainly is completely unequalled in modern times.

higher than the others. In the long run, they are all needed for the continued prospering of mathematics.

Ramanujan's particular talent will seem to be primarily of an algebraic and combinatorial nature. He developed it, for a long time in complete isolation really without any contact with other mathematicians. He had, on his own, acquired an extraordinary skill of manipulation of algorithms, series, continued fractions and so forth, which certainly is completely unequalled in modern times. He seems also to have often taken a very particular delight in the special rather than the general, and if one looks in his *Notebooks*, and also in his letters to Hardy, for instance, he will often state or choose to state a special case or special cases that seem particularly striking, where very clearly he has much more general results underlying them.

He might very well have become a theory-builder, if he had a different and more conventional start and training as a mathematician. Even then, in what has been left in his work, there seems quite clear evidence that he had developed, on his own, a theory of modular forms and equations, for instance, but the precise form of this theory has to be guessed from the isolated results he wrote down in the *Notebooks*. What we have from him are mostly his many and often mysterious results and assertions in the *Notebooks*, and we have his published papers.

It is interesting to see how his published work has been assessed over the years. In the early years, it was clearly the joint work with Hardy on the partition function that drew most of the attention. When his *Collected Papers* appeared, and J E Littlewood wrote a review, he singled this paper out. Most of his review is devoted to trying to analyse how this paper arose. No doubt it was an extremely important paper both in itself and for the powerful tool, the circle method, which was for the first time introduced in Analytic Number Theory. But when it comes to Littlewood's account, it seems to me that it has to be quite wrong on several points.[2]

[2] See the Appendix to this talk in *Atle Selberg, Collected Papers*, Vol I, pages 701-706 concerning this and the following comments on the paper on the partition function.

Littlewood, by the way, wrote his review before the final word had been said about the partition function by H Rademacher's paper in 1937, when he found the exact formula. The paper by Hardy and Ramanujan contained surely a result that was very remarkable in itself; since $p(n)$ is an integer, it allows exact computation on it. But it was not an exact formula. It was a formula with an error tending to zero as n grows and therefore $p(n)$ being an integer, one could find the exact value.

If one looks at Ramanujan's first letter to Hardy, there is a statement there which has some relation to his later work on the partition function, namely about the coefficient of the reciprocal of a certain theta series (a power series with square exponents and alternating signs as coefficients). It gives the leading term in what he claims as an approximate expression for the coefficient. If one looks at that expression, one sees that this is the exact analogue of the leading term in the Rademacher formula for $p(n)$ which shows that Ramanujan, in whatever way he had obtained this, had been led to the correct term of that expression.

In the work on the partition function, studying the paper it seems clear to me that it must have been, in a way, Hardy who did not fully trust Ramanujan's insight and intuition, when he chose the other form of the terms in their expression, for a purely technical reason, which one analyses as not very relevant. I think that if Hardy had trusted Ramanujan more, they should have inevitably ended with the Rademacher series. There is little doubt about that.

Littlewood and Hardy were primarily working with hard analysis and they did not have a strong feeling for modular forms and such things; the generating function for the partition function is essentially a modular form, particularly if one puts in an extra factor $x^{-1/24}$ to the power series. This must have been something that came quite naturally to Ramanujan from the beginning. But to Littlewood, in this review, it seems as if it was an afterthought

I think that a felicitous but unproved conjecture may be of much more consequence for mathematics than the proof of many a respectable theorem.

by a particular stroke of genius that happened later in the development. I find this completely misleading. Littlewood was not present when this happened; he was away from Cambridge during most of the years Ramanujan was there. Littlewood's memory also fails him when he refers to the statement that Ramanujan made in this letter to Hardy as referring to the partition function and not another but rather similarly built function.

At a later stage, when Hardy published the book *12 Lectures on Ramanujan*, Louis J Mordell reviewed this book and he questioned Hardy's assessment that Ramanujan was a man whose native talent was equal to that of Euler or Jacobi. Mordell questions whether the expression "native talent" has any real meaning and claims that one should judge a mathematician by what he has actually done, by which Mordell seems to mean, the theorems he has proved.

By the way, I should say Mordell clearly at no stage seems to have had access to or seen Ramanujan's *Notebooks*. Mordell's assessment seems quite wrong to me. I think that a felicitous but unproved conjecture may be of much more consequence for mathematics than the proof of many a respectable theorem.

Ramanujan's recognition of the multiplicative properties of the coefficients of modular forms that we now refer to as cusp forms and his conjectures formulated in this connection, and their later generalization, have come to play a more central role in the mathematics of today, serving as a kind of focus for the attention of quite a large group of the best mathematicians of our time. Other discoveries like the mock-theta functions are only in the very early stages of being understood and no one can yet assess their real importance. So the final verdict is certainly not in, and it may not be in for a long time, but the estimates of Ramanujan's stature in mathematics certainly have been growing over the years. There is no doubt about that.

It seems quite clear that Hardy's assessment of the degree of Ramanujan's gifts was quite correct, in spite of the great difference that there was between Hardy and Ramanujan in their particular mathematical leanings. One might speculate, although it may be somewhat futile, about what would have happened if Ramanujan had come in contact not with Hardy but with a great mathematician of more similar talents, someone who was more inclined in the algebraic directions, for instance, E Hecke in Germany. This might have perhaps proved much more beneficial and brought out new things in Ramanujan that did not come to fruition by his contact with Hardy. But Hardy deserves greatest credit for recognizing Ramanujan's originality and assisting him and his work in the best way he could.

It seems quite clear that Hardy's assessment of the degree of Ramanujan's gifts was quite correct, in spite of the great difference that there was between Hardy and Ramanujan in their particular mathematical leanings.

I mentioned the assessment by Hardy of Ramanujan and I might quote from his preface to the *Collected Papers*. "Opinions may differ as to the importance of Ramanujan's work, the kind of standard by which it should be judged and the influence which it is likely to have on the mathematics of the future. It has not the simplicity and the inevitableness of the very greatest work; it would be greater if it were less strange. One gift it shows which no one can deny – profound and invincible originality." I think that is a very good phrase actually. "He would probably have been a greater mathematican if he could have been caught and tamed a little in his youth ..." and so forth, wrote Hardy. It is a true statement in one way and in some way it is not. I do not think that Hardy fully understood how the interest for Ramanujan's work would be growing, when he speaks of the influence which it is likely to have on the mathematics of the future. It seems rather clear that he underestimated that. Later developments have certainly shown him wrong on that point.

Hardy refers to Ramanujan also as one of the few romantic figures in the history of mathematics and also, in another place, as the one great romance of his life. Other romantic figures one may think of are, for instance, Galois or Abel. They died of course even younger but they did, although they had their

Other romantic figures one may think of are, for instance Galois or Abel...One may say that none of them seems to have been quite so exclusively committed to mathematics, with no other visible interest in their life as Ramanujan was, and in this, I think, he truly stands completely alone among the mathematicians.

difficulties, come from a somewhat more fortunate environment. One may say that none of them seems to have been quite so exclusively committed to mathematics, with no other visible interest in their life as Ramanujan was, and in this, I think, he truly stands completely alone among the mathematicians. At least from what I have read about him, there was no other real interest for him than mathematics.

It is also futile to speculate on what Ramanujan or, for that matter, Abel or Galois might have accomplished further, if their lives had not been cut short. The peak of a mathematician's work, I think, usually occurs between 30 and 40 years of age. One might have expected in all these cases that the added contribution might have equalled or outweighed their previous work. What is not futile perhaps is to consider what would have been lost if Ramanujan had not finally been given his chance. He is not the first exceptionally gifted person who has suffered rejection in a complexity inflexible and rigid educational system. Galois also failed in his examination. Abel was considered a very mediocre student, also in mathematics, until he got an exceptional mathematics teacher in school who did recognize his talents. Einstein had considerable difficulties in school. And in quite another field, one could mention Thomas Edison who was described by his teachers as having an addled brain.

One sometimes hears it said that true genius will always make its way and be recognized in the end. I do not think that is true.

There really must have been cases in the past where the native talent was finally completely thwarted by an inelastic system and teachers without sufficient understanding for the rare and unusual student for whom an exception should have been made. And I think this lack of consideration could continue usually with impunity because the world would never know what had been lost. The most important lesson that one could draw from Ramanujan's story about the educational system is that allowances should be made for the unusual and perhaps lopsidedly

gifted child with very strong interests in one direction, at all stages of the educational system.

There is also another thing which, I think, is rather important, and this is the school mathematics. I have talked with many others who became mathematicians, about the mathematics they learned in school. Most of them were not particularly inspired by it but started reading on their own, outside of school by some accident or the other, as I myself did. I think mathematics in the school definitely should be revised in such a way that it gives more of a sense of discovery and excitement. I think, in this, the teaching of mathematics often differs from the teaching of the other sciences which in the school is usually better carried out and does give a sense of discovery and excitement. And besides the schools, I think, it is also important for the development of possible future Ramanujans, that public libraries should stock a reasonable amount of mathematical books that could inspire and really interest someone who wants to find something outside of his school curriculum. This is one important thing that can be done in the future for making it easier for any future Ramanujan.

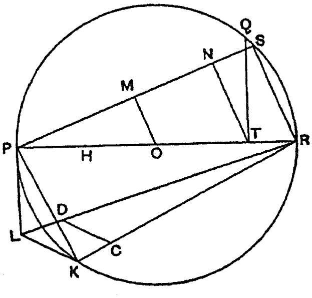
The most important lesson that one could draw from Ramanujan's story about the educational system is that allowances should be made for the unusual and perhaps lopsidedly gifted child with very strong interests in one direction, at all stages of the educational system.

Squaring The Circle, S Ramanujan

Let PQR be a circle with centre O, of which a diameter is PR. Bisect PO at H and let T be the point of trisection of OR nearer R. Draw TQ perpendicular to PR and place the chors $RS = TQ$.

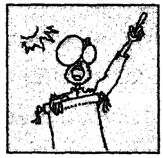
Join PS, and draw OM and TN parallel to RS. Place a chord $PK = PM$, and draw the tangent $PL = MN$. Join RL, RK and KL. Cut off $RC = RH$. Draw CD parallel to KL, meeting RL at D.

Then the square on RD will be equal to the circle PQR approximately.



Journal of the Indian Mathematical Society, V, 1913, 132

Information and Announcements



National Conference on Energy Crisis and Environment

Loyola Institute of Frontier Energy (LIFE) is organizing a National Conference on "Energy Crisis and Environment" on March 7 and 8, 1997. The areas of focus are: Non-Conventional Energy, Environmental Energetics, Environmental Biotechnology, Environmental Engineering, Environmental Chemistry, Environmental Management, and Environmental Ethics. Abstracts (200 words) are welcome before December 15, 96.

For further information write to

Dr Francis P Xavier, S J
Convenor/Director
Loyola Institute of Frontier Energy (LIFE)
Loyola College
Chennai 600034

Errata

Resonance, Vol. 1, No. 10 (1996)

Pages 76-77: In the equations for momentum and energy conservation the parameter β *should be read as* β'*, in the terms involving* M'.

Acknowledgements

Resonance gratefully acknowledges the help received from the following individuals:

Sumanta Baruah	C R Praneshachar
Milind Kolatkar	Jayant Rao
M Murakami	B J Venkatachala

Annual Subscription Rates

	Personal	Institutional
India	Rs 100	Rs 200
Third World Countries	US $ 25	US $ 50
Other Countries	US $ 50 (Students $ 25)	US $ 100

Send your subscriptions by DD or MO in favour of "Indian Academy of Sciences" to Circulation Department, Indian Academy of Sciences, C V Raman Avenue, PB No. 8005, Bangalore 560080, India.

Edited and published by V K Gaur for the Indian Academy of Sciences, Bangalore 560 080. Printed at Thomson Press (I) Ltd., Faridabad 121007.

Guidelines for Authors

Resonance - *journal of science education* is primarily targeted to undergraduate students and teachers. The journal invites contributions in various branches of science and emphasizes a lucid style that will attract readers from diverse backgrounds. A helpful general rule is that at least the first one third of the article should be readily understood by a general audience.

Articles on topics in the undergraduate curriculum, especially those which students often consider difficult to understand, new classroom experiments, emerging techniques and ideas and innovative procedures for teaching specific concepts are particularly welcome. The submitted contributions should not have appeared elsewhere.

Manuscripts should be submitted in *duplicate* to any of the editors. Authors having access to a PC are encouraged to submit an ASCII version on a floppy diskette. If necessary the editors may edit the manuscript substantially in order to maintain uniformity of presentation and to enhance readability. Illustrations and other material if reproduced, must be properly credited; it is the author's responsibility to obtain permission of reproduction (copies of letters of permission should be sent). In case of difficulty, please contact the editors.

Title Authors are encouraged to provide a 4-7 word title and a 4-10 word sub-title. One of these should be a precise technical description of the contents of the article, while the other must attract the general readers' attention.

Author(s) The author's name and mailing address should be provided. A photograph and a brief (in less than 100 words) biographical sketch may be added. Inclusion of phone and fax numbers and e-mail address would help in expediting the processing of manuscripts.

Summary and Brief Provide a 2 to 4 sentence summary, and preferably a one sentence brief for the contents page.

Style and Contents Use simple English. Keep the sentences short. Break up the text into logical units, with readily understandable headings for each. Do not use multiple sub sections. Articles should generally be 1000-2000 words long.

Illustrations Use figures, charts and schemes liberally. A few colour illustrations may be useful. Try to use good quality computer generated images, with neatly labelled axes, clear labels, fonts and shades. Figure captions must be written with care and in some detail. Key features of the illustration may be pointed out in the caption.

Boxes Highlights, summaries, biographical and historical notes and margin notes presented at a level different from the main body of the text and which nevertheless enhance the interest of the main theme can be placed as boxed items. These would be printed in a different typeface. Such a boxed item should fit in a printed page and not exceed 250 words.

Suggested Reading Avoid technical references. If some citations are necessary, mention these as part of the text. A list of suggested readings may be included at the end.

Layout It is preferable to place all the boxes, illustrations and their captions after the main text of the article. The suggested location of the boxes and figures in the printed version may be marked in the text. In the printed version, the main text will occupy two-thirds of each page. The remaining large margin space will be used to highlight the contents of key paragraphs, for figure captions, or perhaps even for small figures. The space is to be used imaginatively to draw attention to the article. Although the editors will attempt to prepare these entries, authors are encouraged to make suitable suggestions and provide them as an annexure.

Book Reviews

The following types of books will be reviewed : (1) text books in subjects of interest to the journal; (2) general books in science brought to the attention of students/teachers; (3) well-known classics; (4) books on educational methods. Books reviewed should generally be affordable to students/teachers (price range Rs.50 to 300).

New books will get preference in review. A list of books received by the academy office will be circulated among the editors who will then decide which ones are to be listed and which to be reviewed.

"In South India today, everyone has heard of Ramanujan. College professors and bicycle rickshaw drivers alike know his story, at least in sketchy outline, just as everyone in the West knows of Einstein. Few can say much about his work, and yet something in the story of his struggle for the chance to pursue his work on his own terms compels the imagination, leaving Ramanujan a symbol for genius, for the obstacle it faces, for the burden it bears, for the pleasure it takes in its own existence".

Robert Kanigel in "The man who knew infinity"

Srinivasa Ramanujan

(December 1887 - April 1920)